

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	3655.8	100.0	3657	6	AR492578	Sequence
2	3654.2	99.9	5590	6	AR136524	Sequence
3	3654.2	99.9	5590	6	CO881401	Sequence
4	3654.2	99.9	5590	6	AR482321	Sequence
5	3654.2	99.9	5590	8	HSU73936	Sequence
6	3654.2	99.9	5596	6	CO873974	Sequence
7	3654.2	99.9	5596	6	CO894703	Sequence
8	3654.2	99.9	5596	6	CS081309	Sequence
9	3654.2	99.9	5596	6	CS024278	Sequence
10	3654.2	99.9	5596	6	AX578038	Sequence
11	3651	99.9	4208	6	AR182312	Sequence
12	3650	99.9	5457	8	AF028593	Homo-sapi
13	3651	99.9	5597	6	CO727152	Sequence
14	3649.4	99.8	4855	6	AR482475	Sequence
15	3649.4	99.8	4855	6	AX375031	Sequence
16	3649.4	99.8	4855	8	HSU61276	Human trans
17	3649.4	99.8	5342	8	AF003837	Homo sapi
18	3647.8	99.8	4208	6	AR169937	Sequence

QY	241	GAGTATCAGTCCCGCGTCACGGCCGGGGGCGCCTTGACAGCTTCGGCTCAGGGTCCA	CGCCT	300
Db	241	GAGTATCAGTCCCGCGTCA	GGCCGGGGGGCCCTGACAGCTTCGGCTCAGGGTCCA	CGCCT
QY	301	GTCAATCGGGGGCAACACCTTCAACCTCAAGGCCAGCGCGGCAACAGCCAA	CCGCATC	360
Db	301	GTCAATCGGGGGCAACACCTTCAACCTCAAGGCCAGCGCGGCAACAGCCAA	CCGCATC	360
QY	361	GTGCTGCTTTTCACTTTTCCGCTGGCCGAGTCTTATACGTTTGTGTGGAGGCGTGGGAT		420
Db	361	GTGCTGCTTTTCACTTTTCCGCTGGCCGAGTCTTATACGTTTGTGTGGAGGCGTGGGAT		420
QY	421	TCCAGTAATGACACCGTTCAACCTGACAGTATATTTGAAAAGGCTTCTCACTCGGGCATG		480
Db	421	TCCAGTAATGACACCGTTCAACCTGACAGTATATTTGAAAAGGCTTCTCACTCGGGCATG		480
QY	481	ATCAACCCCGAGCCGGCAGTGGCAGACGCTGAAGCAGAACACGGGCGTTGCCACTTTGAG		540
Db	481	ATCAACCCCGAGCCGGCAGTGGCAGACGCTGAAGCAGAACACGGGCGTTGCCACTTTGAG		540
QY	541	TATCAGATCCGCGTGA	CGTGTGATGACTATCTATGCGCTTTGGCTGYAATAAGTCTTGC	600
Db	541	TATCAGATCCGCGTGA	CGTGTGATGACTATCTATGCGCTTTGGCTGYAATAAGTCTTGC	600
QY	601	CGCCCCAGAGTACTTTTGGACATATGCTGTCACCAAGATGCAACAAACTTGC		660
Db	601	CGCCCCAGAGTACTTTTGGACATATGCTGTCACCAAGATGCAACAAACTTGC		660
QY	661	ATGGAAGCTGGATGGGCCCCGAATGTAAACAGAGCTATTTGCCACAAAGCTGCAGTCC		720
Db	661	ATGGAAGCTGGATGGGCCCCGAATGTAAACAGAGCTATTTGCCACAAAGCTGCAGTCC		720
QY	721	AAGCATGGGCTTGGAACTCCCAAGTGA	CTGACAGTGCACAGTAYGGCTGGCAAGGCGCTG	780
Db	721	AAGCATGGGCTTGGAACTCCCAAGTGA	CTGACAGTGCACAGTAYGGCTGGCAAGGCGCTG	780
QY	781	TACTGTGATAGTGCATCCACACCCGGGATCGCTCCACGGCATCTGTAATGAGCCCTGG		840
Db	781	TACTGTGATAGTGCATCCACACCCGGGATCGCTCCACGGCATCTGTAATGAGCCCTGG		840
QY	841	CAGTGCCTCTGTGAGACCAACTGGGGGGCGGACAGCTCTGTGACAAAGATCTCAATTA	CTGT	900
Db	841	CAGTGCCTCTGTGAGACCAACTGGGGGGCGGACAGCTCTGTGACAAAGATCTCAATTA	CTGT	900
QY	901	GGGACTCATAGCCGTGTCTCAACGGGGGAACTTGTAGCAACACAGGGCCCTGACAAATAT		960
Db	901	GGGACTCATAGCCGTGTCTCAACGGGGGAACTTGTAGCAACACAGGGCCCTGACAAATAT		960
QY	961	CAGTGTTCCTGCTGAGGGGTATT	CAGGACCCCACTGTGAAATTTGCTGAGCAGCCCTGC	1020
Db	961	CAGTGTTCCTGCTGAGGGGTATT	CAGGACCCCACTGTGAAATTTGCTGAGCAGCCCTGC	1020
QY	1021	CTCTCTGATCCCTGTCAACACAGAGGACAGTGTAAAGGAGACCTCCCTGGGCTTTGAGTGT		1080
Db	1021	CTCTCTGATCCCTGTCAACACAGAGGACAGTGTAAAGGAGACCTCCCTGGGCTTTGAGTGT		1080
QY	1081	GAGTGTTCCTGAGGCTGACACGGGCCCAATGCTCTACAAACATGATGATGTTCTCTCT		1140
Db	1081	GAGTGTTCCTGAGGCTGACACGGGCCCAATGCTCTACAAACATGATGATGTTCTCTCT		1140
QY	1141	AATACTGTTCCACGGGGGACCTGCCAGACCTGGTTAACGGATTTTAAAGTGTGTGTC		1200
Db	1141	AATACTGTTCCACGGGGGACCTGCCAGACCTGGTTAACGGATTTTAAAGTGTGTGTC		1200
QY	1201	CCCCCAGAGTGGACTGGGAAACAGTGCAGTGTAGATGCAATGAATGTGAGGCCAAACCT		1260
Db	1201	CCCCCAGAGTGGACTGGGAAACAGTGCAGTGTAGATGCAATGAATGTGAGGCCAAACCT		1260
QY	1261	TGTGTAAACGGCCAAATCTCTGTAAAGAAATCTATTCAGCTACTACTGCGACTGTCTTCCC		1320
Db	1261	TGTGTAAACGGCCAAATCTCTGTAAAGAAATCTATTCAGCTACTACTGCGACTGTCTTCCC		1320
QY	1321	GGCTGGATGGGTACAGAAATTTGTGACATAAATATTAATATGACTGCTTGGCCAGTGTGCAGAA	T	1380
Db	1321	GGCTGGATGGGTACAGAAATTTGTGACATAAATATTAATATGACTGCTTGGCCAGTGTGCAGAA	T	1380
QY	1381	GACGGCTCCTGTGCGGATTTGGTAAATGGTATCCCTGTACTCTGTACTCTGACCTGGGGGT		1440
Db	1381	GACGGCTCCTGTGCGGATTTGGTAAATGGTATCCCTGTACTCTGTACTCTGACCTGGGGGT		1440
QY	1441	GCGGATCACTGTGAGAGAGACATCCATGAATGTGCGCAGCAACCCCTGTTTGAATGGGGGT		1500
Db	1441	GCGGATCACTGTGAGAGAGACATCCATGAATGTGCGCAGCAACCCCTGTTTGAATGGGGGT		1500
QY	1501	CACTGTGAGAAATCAACAGATTCACAGTCTGTGTGCCACCTGTGTTTCTCTGGAAC		1560
Db	1501	CACTGTGAGAAATCAACAGATTCACAGTCTGTGTGCCACCTGTGTTTCTCTGGAAC		1560
QY	1561	CTTGTGAGCTGGACATCGATTTATTTGTAGGCTTAATCCCTGCCAGAACGGTGGCCAGTGC		1620
Db	1561	CTTGTGAGCTGGACATCGATTTATTTGTAGGCTTAATCCCTGCCAGAACGGTGGCCAGTGC		1620
QY	1621	TACAAACCGTCCAGTGA	CTATTTCTGCAAGTCCCGGAGCACTATGAGGGCAAGAACTGC	1680
Db	1621	TACAAACCGTCCAGTGA	CTATTTCTGCAAGTCCCGGAGCACTATGAGGGCAAGAACTGC	1680
QY	1681	TCACACCTGAAAGACCACTGCCGACGACCCCTCTGTGAAGTGAATGACAGCTGCACAGTG		1740
Db	1681	TCACACCTGAAAGACCACTGCCGACGACCCCTCTGTGAAGTGAATGACAGCTGCACAGTG		1740
QY	1741	GCCATGGCTTCCAAACGACACCTGAAAGGGGTGCGGTATATTTCTCCAAACCTGTGTGTGT		1800
Db	1741	GCCATGGCTTCCAAACGACACCTGAAAGGGGTGCGGTATATTTCTCCAAACCTGTGTGTGT		1800
QY	1801	CCTCAACGGGAGTGCAGAGTCACTGGGAGGCAAAATTCACCTGTGACTGTAAACAAAGGC		1860
Db	1801	CCTCAACGGGAGTGCAGAGTCACTGGGAGGCAAAATTCACCTGTGACTGTAAACAAAGGC		1860
QY	1861	TTCAACGGGAAACATCTGCATGAAATATTAATATGACTGTGAGAGCAACCTTTGTGAAAC		1920
Db	1861	TTCAACGGGAAACATCTGCATGAAATATTAATATGACTGTGAGAGCAACCTTTGTGAAAC		1920
QY	1921	GTTGGACATTCGATCGATGGTCACTCTACAGTGCATCTGTAGTCAACGGCTGGGAG		1980
Db	1921	GTTGGACATTCGATCGATGGTCACTCTACAGTGCATCTGTAGTCAACGGCTGGGAG		1980
QY	1981	GGGGCTCTCTGTGAAACCAATATTAATGACTGCAGCGCAACCCCTGCCACAATGGGGGC		2040
Db	1981	GGGGCTCTCTGTGAAACCAATATTAATGACTGCAGCGCAACCCCTGCCACAATGGGGGC		2040
QY	2041	ACGTGTGCGGACCTGTGTCATCACTTCTACTGTGACTGTAAATAATGGGTGGAAAGAAAG		2100
Db	2041	ACGTGTGCGGACCTGTGTCATCACTTCTACTGTGACTGTAAATAATGGGTGGAAAGAAAG		2100
QY	2101	ACTGCGACCTCACTGTCAGTGCAGTGTGATGAGGCCCACTGTGCAACAAACGGTGGCACTGC		2160
Db	2101	ACTGCGACCTCACTGTCAGTGCAGTGTGATGAGGCCCACTGTGCAACAAACGGTGGCACTGC		2160
QY	2161	TATGATGAGGGGGAATGCTTTTAAAGTGCATGTGTCTGCGGCTGGGAAGGAAACAACTGT		2220
Db	2161	TATGATGAGGGGGAATGCTTTTAAAGTGCATGTGTCTGCGGCTGGGAAGGAAACAACTGT		2220
QY	2221	AACATAGCCCGAAACAGTAGTGCCTGCCAACCCCTGCCATAATGGGGGSCACATGTGTG		2280
Db	2221	AACATAGCCCGAAACAGTAGTGCCTGCCAACCCCTGCCATAATGGGGGSCACATGTGTG		2280
QY	2281	GTCAACGGCGAGTCTTTTACAGTGGTCTGCAAGGAAAGGCTGGAGGGGGCCCATCTGTGCT		2340
Db	2281	GTCAACGGCGAGTCTTTTACAGTGGTCTGCAAGGAAAGGCTGGAGGGGGCCCATCTGTGCT		2340
QY	2341	CAGNAATACCAATGACTGCGACCCCTCATCTCTGTGTACAAAGCGGCACCTGTGTGGATGGA		2400
Db	2341	CAGNAATACCAATGACTGCGACCCCTCATCTCTGTGTACAAAGCGGCACCTGTGTGGATGGA		2400
QY	2401	GACAACTGGTACCGGTGCGGAATGTGCCCGCGGGTTTTTCTGTGGGCGCGACTGCAAGAAATAAC		2460

```
Db 2401 GACAACTGGTACCGGTGCGAATGTGCCCCGGGTTTTGCTGGGCCCGGACTGCAGAAATAAC 2460
Qy 2461 ATCAATGAATGCCAGTCTTTCACCTTGTGCTTTGGAGCGACCTGTGTGGATGAGATCAAT 2520
Db 2461 ATCAATGAATGCCAGTCTTTCACCTTGTGCTTTGGAGCGACCTGTGTGGATGAGATCAAT 2520
Qy 2521 GGCTACCGGTGTGTCTGCCCTCCAGGCGACAGTGTGTGCCAAGTGCACGAAAGTTTCAGGG 2580
Db 2521 GGCTACCGGTGTGTCTGCCCTCCAGGCGCACAGTGTGTGCCAAGTGCACGAAAGTTTCAGGG 2580
Qy 2581 AGACCTTGATCAACATGGGGAGTGTGATACAGATGGGGCCCAAAATGGGATGATGACTGT 2640
Db 2581 AGACCTTGATCAACATGGGGAGTGTGATACAGATGGGGCCCAAAATGGGATGATGACTGT 2640
Qy 2641 AATACCTGCCAGTGCCTTGAATGGAACGATGCCCTGTCTCAAGGTCTGTGTGGCCCTCGA 2700
Db 2641 AATACCTGCCAGTGCCTTGAATGGAACGATGCCCTGTCTCAAGGTCTGTGTGGCCCTCGA 2700
Qy 2701 CCTTGCTGTCTCAAAAGGGGCAAGCGAGTGGCCCAAGCGGCGAGAGCTGCATCCCAATC 2760
Db 2701 CCTTGCTGTCTCAAAAGGGGCAAGCGAGTGGCCCAAGCGGCGAGAGCTGCATCCCAATC 2760
Qy 2761 CTGGAGCAGAGTCTTCCCTCCACCCCTGCACTGTGTGGGCGAGTGTGCTTCCAGT 2820
Db 2761 CTGGAGCAGAGTCTTCCCTCCACCCCTGCACTGTGTGGGCGAGTGTGCTTCCAGT 2820
Qy 2821 CTCAGCGGTGAAGCAAAAGTGCACCTCTGACTCTTATACAGAGTAATCTGTGCGAAC 2880
Db 2821 CTCAGCGGTGAAGCAAAAGTGCACCTCTGACTCTTATACAGAGTAATCTGTGCGAAC 2880
Qy 2881 ATCAATATTAACCTTTAAAGGAGATGATGTACAGAGTCTTACTACGAGACATTTGC 2940
Db 2881 ATCAATATTAACCTTTAAAGGAGATGATGTACAGAGTCTTACTACGAGACATTTGC 2940
Qy 2941 AGTGAATTTGAGGAATTTGAATATTTTGAAGAAATTTTCCGCTGAATATTCATCTACATC 3000
Db 2941 AGTGAATTTGAGGAATTTGAATATTTTGAAGAAATTTTCCGCTGAATATTCATCTACATC 3000
Qy 3001 GCTTTCAGAGCTTCCCTTCAGCGAACAAATGAATACATGTGGCCATTTCTGCTGAAGAT 3060
Db 3001 GCTTTCAGAGCTTCCCTTCAGCGAACAAATGAATACATGTGGCCATTTCTGCTGAAGAT 3060
Qy 3061 ATACGGGATGATGGGAACCCGATCAAGGAAATCACTGACAAAAATAATCGATCTTTTGTAGT 3120
Db 3061 ATACGGGATGATGGGAACCCGATCAAGGAAATCACTGACAAAAATAATCGATCTTTTGTAGT 3120
Qy 3121 AAACGTGATGGAACACAGTCTGCTGATTTGCTGCGGTTCGAGAGTTCAGAGCGG 3180
Db 3121 AAACGTGATGGAACACAGTCTGCTGATTTGCTGCGGTTCGAGAGTTCAGAGCGG 3180
Qy 3181 CCTCTGAAGAACACAGATTTCTTGTTCCTTGTCTGAGCTCTGCTTAACTGTGGCT 3240
Db 3181 CCTCTGAAGAACACAGATTTCTTGTTCCTTGTCTGAGCTCTGCTTAACTGTGGCT 3240
Qy 3241 TGGATCTGTTGTGTGTGACGGCTTCTACTGTGTGCTGCGGAGCGGCGGAGCCGGGC 3300
Db 3241 TGGATCTGTTGTGTGTGACGGCTTCTACTGTGTGCTGCGGAGCGGCGGAGCCGGGC 3300
Qy 3301 AGCCACACACACTCAGCTCTGAGGACAAACACACAAAGAGTGTGGGAGAGCTGAAAC 3360
Db 3301 AGCCACACACACTCAGCTCTGAGGACAAACACCAACCAACAGTGTGGGAGAGCTGAAAC 3360
Qy 3361 CAGATCAAAAACCCATTGAGAAACATGGGGCCCAACAGCGTCCCATCAAGGATTAAGAG 3420
Db 3361 CAGATCAAAAACCCATTGAGAAACATGGGGCCCAACAGCGTCCCATCAAGGATTAAGAG 3420
Qy 3421 AACAGAACTCCAAATGTCTTAAATTAAGGACACAAATTTGAAAGTGAAGGAGGACGAC 3480
Db 3421 AACAGAACTCCAAATGTCTTAAATTAAGGACACAAATTTGAAAGTGAAGGAGGACGAC 3480
Qy 3481 ATGACAAAACACAGCAGAAAGCCCGGTTTTGGCAGCAGCGCGGCTTATGCTGGTAGAC 3540
```

```
Db 3481 ATGACAAAACACAGCAGAAAGCCCGGTTTTGGCAAGCAGCGCGGTATACGCTGGTAGAC 3540
Qy 3541 AGAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600
Db 3541 AGAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600
Qy 3601 GACACAGAGACTTGGAAAGTGGCCAGAGCTTAAACCGAATGGAGTACATCGTATAG 3657
Db 3601 GACACAGAGACTTGGAAAGTGGCCAGAGCTTAAACCGAATGGAGTACATCGTATAG 3657

RESULT 2
AR136524
LOCUS AR136524 5590 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 1 from patent US 6136952.
ACCESSION AR136524
VERSION AR136524.1 GI:14477196
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 5590)
TITLE Li, L. and Hood, L.
JOURNAL Human jagged polypeptide, encoding nucleic acids and methods of use
FEATURES Patent: US 6136952-A 1 24-OCT-2000;
Location/Qualifiers
source 1..5590
/mol_type="unknown"
/organism="unassigned DNA"

Query Match 99.9%; Score 3654.2; DB 6; Length 5590;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3653; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATCGTGTCCCGACGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Db 414 ATCGTGTCCCGACGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 473
Qy 61 CTCTGTGCTCCCGACGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 474 CTCTGTGCTCCCGACGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 533
Qy 121 TCCATGACAGAACCTGTAAACCGGAGAGTGCAGAAACCGGAACTGTCTGCGCGCGCGCGCG 180
Db 534 TCCATGACAGAACCTGTAAACCGGAGAGTGCAGAAACCGGAACTGTCTGCGCGCGCGCGCG 593
Qy 181 CCGGAGACCGCAAGTGCACCCCGGACGAGTGTGACACATCTTCAAGTGTGCGCTCAAG 240
Db 594 CCGGAGACCGCAAGTGCACCCCGGACGAGTGTGACACATCTTCAAGTGTGCGCTCAAG 653
Qy 241 GAGTATCAGTCCCGCGCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 300
Db 654 GAGTATCAGTCCCGCGCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 713
Qy 301 GTCATCGGGGCAACACCTTCAACCTCAAGCGCAGCGCGCGCGCGCGCGCGCGCGCGCATC 360
Db 714 GTCATCGGGGCAACACCTTCAACCTCAAGCGCAGCGCGCGCGCGCGCGCGCGCGCATC 773
Qy 361 GTGCTGCTTTTCAATTTGCGCTGCGCGAGTGTCTATACGTTGCTTGGAGGCGTGGGAT 420
Db 774 GTGCTGCTTTTCAATTTGCGCTGCGCGAGTGTCTATACGTTGCTTGGAGGCGTGGGAT 833
Qy 421 TCCAGTAATGACACCGTTCAACCTGACAGTATTTATGAAAAGGCTTCTCACTCGGGCATG 480
Db 834 TCCAGTAATGACACCGTTCAACCTGACAGTATTTATGAAAAGGCTTCTCACTCGGGCATG 893
Qy 481 ATCAACCCAGCGCGAGTGTGCGAGACGCTCAAGCAGAAACCGGCGGTGTGCCACTTTGAG 540
Db 894 ATCAACCCAGCGCGAGTGTGCGAGACGCTCAAGCAGAAACCGGCGGTGTGCCACTTTGAG 953
Qy 541 TATCAGATCCGCTGACCTGTGATGACTACTATATGCTTGGCTTGGCTTGAATAGTTCTGC 600
```

Db	954	TATCAGATCCGCGTGACCTGTGATGACTACTATATGGCTTTTGGCTGCAATAAGTTCTGC	1013
Qy	601	CGCCCCAGAGATGACTTTCTTTGGACACTATGCTCTGTGACACAGAAATGGCAACAAACTTGC	660
Db	1014	CGCCCCAGAGATGACTTTCTTTGGACACTATGCTCTGTGACACAGAAATGGCAACAAACTTGC	1073
Qy	661	ATGGAAGGCTGGATGGSCCCGGAATGTAACAGAGCTATTTTGCCGACAAAGGCTCGAGTCCT	720
Db	1074	ATGGAAGGCTGGATGGSCCCGGAATGTAACAGAGCTATTTTGCCGACAAAGGCTCGAGTCCT	1133
Qy	721	AAGCATGGGCTTTGCCAACTCCAGGTGACTGCAAGTGGCAGTAYGGCTGGCAAGGCTG	780
Db	1134	AAGCATGGGCTTTGCCAACTCCAGGTGACTGCAAGTGGCAGTAYGGCTGGCAAGGCTG	1193
Qy	781	TACTGTGATTAAGTGATCCACACACCCGGGATGGGTCAACGGGATCTGTATATGAGCCCTGG	840
Db	1194	TACTGTGATTAAGTGATCCACACACCCGGGATGGGTCAACGGGATCTGTATATGAGCCCTGG	1253
Qy	841	CAGTGCCCTCTGTGAGACCAACTGGGGCGGCAGCTCTGTGACAAAGATCTCAATTAATCTGT	900
Db	1254	CAGTGCCCTCTGTGAGACCAACTGGGGCGGCAGCTCTGTGACAAAGATCTCAATTAATCTGT	1313
Qy	901	GGGACTCATCAGCCGCTGTCTCAACGGGGGAACTTTGTAGCAACACAGGCCCTGCAAAATAT	960
Db	1314	GGGACTCATCAGCCGCTGTCTCAACGGGGGAACTTTGTAGCAACACAGGCCCTGCAAAATAT	1373
Qy	961	CAGTGCTCTGCTGAGAGCAACTGGGGCGGCAGCTCTGTGACAAAGATCTCAATTAATCTGT	1020
Db	1374	CAGTGCTCTGCTGAGAGCAACTGGGGCGGCAGCTCTGTGACAAAGATCTCAATTAATCTGT	1433
Qy	1021	CTCTCTCATCCCTGTCAACACAGAGGAGCTGTAAAGGAGACCTCCCTGGGCTTTGAGTGT	1080
Db	1434	CTCTCTCATCCCTGTCAACACAGAGGAGCTGTAAAGGAGACCTCCCTGGGCTTTGAGTGT	1493
Qy	1081	GAGTGTTCGCCAGGCTGGACCGGCCCCACATGCTCTACAAACATATGATGACTGTTCCTCT	1140
Db	1494	GAGTGTTCGCCAGGCTGGACCGGCCCCACATGCTCTACAAACATATGATGACTGTTCCTCT	1553
Qy	1141	AATAACTGTTCCACCGGGGACCTGCGCAGGACCTGCTGTTAAACGATTTAAGTGTGTGTC	1200
Db	1554	AATAACTGTTCCACCGGGGACCTGCGCAGGACCTGCTGTTAAACGATTTAAGTGTGTGTC	1613
Qy	1201	CCCCACAGTGGACTGGGAAAACCTGCGAGTTAGATGCAAAATGATGAGGCGCAACCT	1260
Db	1614	CCCCACAGTGGACTGGGAAAACCTGCGAGTTAGATGCAAAATGATGAGGCGCAACCT	1673
Qy	1261	TGTTAAACGCCAAATCCTGTAAGAACTCTCATTGCCAGCTACTCTGCGACTGCTTTCCC	1320
Db	1674	TGTTAAACGCCAAATCCTGTAAGAACTCTCATTGCCAGCTACTCTGCGACTGCTTTCCC	1733
Qy	1321	GGCTGGATGGGTGAGAAATTTGACATAAATTAATGACTGCTGCGGCTTGGCCAGTGTGAGAT	1380
Db	1734	GGCTGGATGGGTGAGAAATTTGACATAAATTAATGACTGCTGCGGCTTGGCCAGTGTGAGAT	1793
Qy	1381	GACCCCTCTGTGCGGATTTGGTTAAATGGTTATCGCTGTATCTGTCACCTGGCTATGCA	1440
Db	1794	GACCCCTCTGTGCGGATTTGGTTAAATGGTTATCGCTGTATCTGTCACCTGGCTATGCA	1853
Qy	1441	GGCGATCACTGTGAGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGGT	1500
Db	1854	GGCGATCACTGTGAGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGGT	1913
Qy	1501	CACGTGAGAAATGAATCAACAGATTCAGTGTCTGTGTCTGCTCCACTGGTTTCTCGGAAAC	1560
Db	1914	CACGTGAGAAATGAATCAACAGATTCAGTGTCTGTGTCTGCTCCACTGGTTTCTCGGAAAC	1973
Qy	1561	CTCTGTGAGTGGACATCGATTTATTTGTAGCCCTTAATCCCTGCCAGAACGGTGGCCAGTGC	1620
Db	1974	CTCTGTGAGTGGACATCGATTTATTTGTAGCCCTTAATCCCTGCCAGAACGGTGGCCAGTGC	2033
Qy	1621	TACAACCGTGCAGTGAATTTCTTCGAAGTGGCCCGGAGGACTATGAGGGCAAGAACTGC	1680
Db	2034	TACAACCGTGCAGTGAATTTCTTCGAAGTGGCCCGGAGGACTATGAGGGCAAGAACTGC	2093

Qy	1681	TCACCTCTGAAGAACCACCTGCGCAGACCCCTGTGAAAGTGATTTCACAGCTGCACAGTG	1740
Db	2094	TCACCTCTGAAGAACCACCTGCGCAGACCCCTGTGAAAGTGATTTCACAGCTGCACAGTG	2153
Qy	1741	GCATGGCTTTCCAAACGACACACCTGAAGGGGTGCGGTATATTTCTCTCAAACGTCGTGGT	1800
Db	2154	GCATGGCTTTCCAAACGACACACCTGAAGGGGTGCGGTATATTTCTCTCAAACGTCGTGGT	2213
Qy	1801	CTTCACGGGAAGTGAAGAGTCAAGTCCGAGGAGCAAAATTCACCTGTGACTGTATCAAAAGC	1860
Db	2214	CTTCACGGGAAGTGAAGAGTCAAGTCCGAGGAGCAAAATTCACCTGTGACTGTATCAAAAGC	2273
Qy	1861	TTTCAACGGGAACATPACTGCCATGAAATATTAATGACTGTGAGAGCAACCTCTGTGAAAC	1920
Db	2274	TTTCAACGGGAACATPACTGCCATGAAATATTAATGACTGTGAGAGCAACCTCTGTGAAAC	2333
Qy	1921	GGTGGCACTTTGCATCGATGGTCAACTCTCTCAAGTGCATCTGTATGACGCTGGAG	1980
Db	2334	GGTGGCACTTTGCATCGATGGTCAACTCTCTCAAGTGCATCTGTATGACGCTGGAG	2393
Qy	1981	GGGGCCCTA CTGTGAAAACCAATTAATGACTGCGAGCCAGAACCCCTGCCCAATGGGGC	2040
Db	2394	GGGGCCCTA CTGTGAAAACCAATTAATGACTGCGAGCCAGAACCCCTGCCCAATGGGGC	2453
Qy	2041	ACGTGTGCGGACCTGTGTCAATGACTTCTACTGTGACTGTAAAAATGGGTGAAAAGGAAAG	2100
Db	2454	ACGTGTGCGGACCTGTGTCAATGACTTCTACTGTGACTGTAAAAATGGGTGAAAAGGAAAG	2513
Qy	2101	ACCTGCCACTCAGCTGACAGTGTGATGAGGCCACGTCGCAACAAACGCTGGGCACTGC	2160
Db	2514	ACCTGCCACTCAGCTGACAGTGTGATGAGGCCACGTCGCAACAAACGCTGGGCACTGC	2573
Qy	2161	TATGATGAGGGGATGCTTTTAAAGTGATGTGTCTTGGGGCTGGGAGGAAACCACTGT	2220
Db	2574	TATGATGAGGGGATGCTTTTAAAGTGATGTGTCTTGGGGCTGGGAGGAAACCACTGT	2633
Qy	2221	AACATAGCCCGAAACAGTAGTCCCTGCCCAACCCCTGCCATTAATGGGGGCAATGTGTG	2280
Db	2634	AACATAGCCCGAAACAGTAGTCCCTGCCCAACCCCTGCCATTAATGGGGGCAATGTGTG	2693
Qy	2281	GTCAAACGGCGAGTCTTTACGTGCTGTGCAAGGAAAGGCTGGGAGGGGCCCATCTGTGCT	2340
Db	2694	GTCAAACGGCGAGTCTTTACGTGCTGTGCAAGGAAAGGCTGGGAGGGGCCCATCTGTGCT	2753
Qy	2341	CAGAATACCAATGACTGCGAGCCCTCATCCCTGTTCACACAGCGGCACCTGTGTGGATGGA	2400
Db	2754	CAGAATACCAATGACTGCGAGCCCTCATCCCTGTTCACACAGCGGCACCTGTGTGGATGGA	2813
Qy	2401	GACAACCTGGTACCGGTGCGAATGTGCCCGGGTTTTGCTGGGCCGCACTGCAGAAATAAAC	2460
Db	2814	GACAACCTGGTACCGGTGCGAATGTGCCCGGGTTTTGCTGGGCCGCACTGCAGAAATAAAC	2873
Qy	2461	ATCAATGAATGCCAGTCTTCACTTTGTGCTTTGGAGCGCACTGTGTGGATGAGATCAAT	2520
Db	2874	ATCAATGAATGCCAGTCTTCACTTTGTGCTTTGGAGCGCACTGTGTGGATGAGATCAAT	2933
Qy	2521	GGCTACCGGTGTGTCTGCCCTTCAAGGCAAGTGGTGGCAGTGGCCAGGAGTTTCAGGG	2580
Db	2934	GGCTACCGGTGTGTCTGCCCTTCAAGGCAAGTGGTGGCAGTGGCCAGGAGTTTCAGGG	2993
Qy	2581	AGACCTTTGCATCACTAGGGGAGTGTATACAGATGGGGCCAAATGGGATGATGACTGT	2640
Db	2994	AGACCTTTGCATCACTAGGGGAGTGTATACAGATGGGGCCAAATGGGATGATGACTGT	3053
Qy	2641	AATACTGCCAGTGCCTGAAATGGAACGATCCCTGTCTCAAAGGTCTGGTGTGGCCCTCGA	2700
Db	3054	AATACTGCCAGTGCCTGAAATGGAACGATCCCTGTCTCAAAGGTCTGGTGTGGCCCTCGA	3113
Qy	2701	CCTTGGCTGTCTCCAAAGGGGCAAGCGAGTGGCCCGGAGGAGCTGCATCCCATC	2760
Db	3114	CCTTGGCTGTCTCCAAAGGGGCAAGCGAGTGGCCCGGAGGAGCTGCATCCCATC	3173

Qy	841	CAGTGCCTCTGTGAGACCAACTGGGGCGGCAGCTCTGTGACAAAGATCTCAATTACTGT	900
Db	1254	CAGTGCCTCTGTGAGACCAACTGGGGCGGCAGCTCTGTGACAAAGATCTCAATTACTGT	1313
Qy	901	GGGACTCATCAGCGGTGTCTCAACGGGGGGAACCTTGTAGCAACACACAGCCCTCGACAAATAT	960
Db	1314	GGGACTCATCAGCGGTGTCTCAACGGGGGGAACCTTGTAGCAACACACAGCCCTCGACAAATAT	1373
Qy	961	CAGTGTTCCTGCGCTGAGGGGTATTCAGGACCCCAACTGTGAAATTCGTGACGACGCTGC	1020
Db	1374	CAGTGTTCCTGCGCTGAGGGGTATTCAGGACCCCAACTGTGAAATTCGTGACGACGCTGC	1433
Qy	1021	CTCTCTGATCCCTGTGACAAACAGAGCGAGCTGTAAAGAGACCTCCCTGGGCTTTTGAAGTGT	1080
Db	1434	CTCTCTGATCCCTGTGACAAACAGAGCGAGCTGTAAAGAGACCTCCCTGGGCTTTTGAAGTGT	1493
Qy	1081	GAGTGTTCCTGAGGCTGGACCGGCCCCACATGCTCTCAAAACATTTGATGACTGTTCCT	1140
Db	1494	GAGTGTTCCTGAGGCTGGACCGGCCCCACATGCTCTCAAAACATTTGATGACTGTTCCT	1553
Qy	1141	AATTAACCTGTTCCACCGGGGCACTGCGAGGACCTGTTTAAACGATTTTAAAGTGTGTGC	1200
Db	1554	AATTAACCTGTTCCACCGGGGCACTGCGAGGACCTGTTTAAACGATTTTAAAGTGTGTGC	1613
Qy	1201	CCCCCAGTGGACTGGGAAAACGTGCCAGTTAGATGCAAAATGAACTGTGAGGCAAAACCT	1260
Db	1614	CCCCCAGTGGACTGGGAAAACGTGCCAGTTAGATGCAAAATGAACTGTGAGGCAAAACCT	1673
Qy	1261	TGTGTAAACGCCAAATCCTGTGAAGATCTCATTTGCCAGCTACTTCCGACTGTCTTCCC	1320
Db	1674	TGTGTAAACGCCAAATCCTGTGAAGATCTCATTTGCCAGCTACTTCCGACTGTCTTCCC	1733
Qy	1321	GGCTGGATGGGTGAGAAATTTGACATATAATTAATGACTGCTTGGCCGACGTGTCAGAAAT	1380
Db	1734	GGCTGGATGGGTGAGAAATTTGACATATAATTAATGACTGCTTGGCCGACGTGTCAGAAAT	1793
Qy	1381	GACGCTTCCTGTCGGGATTTGGTTTAAATGTTATCGCTGTATCTGTCCACCTGGCTATGCA	1440
Db	1794	GACGCTTCCTGTCGGGATTTGGTTTAAATGTTATCGCTGTATCTGTCCACCTGGCTATGCA	1853
Qy	1441	GGCGATCACTGTGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGGT	1500
Db	1854	GGCGATCACTGTGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGGT	1913
Qy	1501	CAGTGTGAGAAATGAACAGATTCAGTGTCTGTGTCCACCTGGCTTCTCGGAAC	1560
Db	1914	CAGTGTGAGAAATGAACAGATTCAGTGTCTGTGTCCACCTGGCTTCTCGGAAC	1973
Qy	1561	CTCTGTGAGTGGACATCGATTATTTGTGAGCTTAATCCCTGCCAGAACGGTGCCAGTGC	1620
Db	1974	CTCTGTGAGTGGACATCGATTATTTGTGAGCTTAATCCCTGCCAGAACGGTGCCAGTGC	2033
Qy	1621	TACAACTGTCAGTGAATTTCTGCAAGTGGCCCGAGGACTATGAGGGCAAGAACTGC	1680
Db	2034	TACAACTGTCAGTGAATTTCTGCAAGTGGCCCGAGGACTATGAGGGCAAGAACTGC	2093
Qy	1681	TCACACTGAAAGACCACTGCGCGACGACCCCTGTGGAAGTGTGACGTGCACAGTG	1740
Db	2094	TCACACTGAAAGACCACTGCGCGACGACCCCTGTGGAAGTGTGACGTGCACAGTG	2153
Qy	1741	GCCATGGCTTCCACGACACACCTGGAAGGGGTGGGTATATTTCCCTCCACGCTGTGGT	1800
Db	2154	GCCATGGCTTCCACGACACACCTGGAAGGGGTGGGTATATTTCCCTCCACGCTGTGGT	2213
Qy	1801	CCTCAGCGGAAGTGCAGAGTCAGTCCGGGAGGCAAAATTCACCTGTGACTGTGTAACAAAGGC	1860
Db	2214	CCTCAGCGGAAGTGCAGAGTCAGTCCGGGAGGCAAAATTCACCTGTGACTGTGTAACAAAGGC	2273
Qy	1861	TTCAACGGGAACATCTGCGCATGAAATTAATTAATGACTGTGAGGACCAACCTTTGAGAAC	1920
Db	2274	TTCAACGGGAACATCTGCGCATGAAATTAATTAATGACTGTGAGGACCAACCTTTGAGAAC	2333
Qy	1921	GGTGGCACTTGCATCGATGGTGTCAACTCTTACAAAGTGCATCTGTAGTGACGGCTGGAG	1980

Db	2334	GGTGGCACTTGCATCGATGGTGTCAACTCTTAACAAGTGCATCTGTAGTGACGGCTGGAG	2393
Qy	1981	GGGGCCTACTGTGAAAACCAATATTAAATGACTGCGAGCCAGAAACCCCTGCCACAAATGGGGC	2040
Db	2394	GGGGCCTACTGTGAAAACCAATATTAAATGACTGCGAGCCAGAAACCCCTGCCACAAATGGGGC	2453
Qy	2041	ACGTGTCGCGACCTGTCAATGACTTCTACTGTGACTGTAAATAATGGGTGGAAGGAAAG	2100
Db	2454	ACGTGTCGCGACCTGTCAATGACTTCTACTGTGACTGTAAATAATGGGTGGAAGGAAAG	2513
Qy	2101	ACCTGCCACTCAGCTGACAGTGTGATGAGGCCACGTGCAACAAACCGTGGCAACTGC	2160
Db	2514	ACCTGCCACTCAGCTGACAGTGTGATGAGGCCACGTGCAACAAACCGTGGCAACTGC	2573
Qy	2161	TATGATGAGGGGATGCTTTTAAAGTGCATGTCTTGGGGCTGGGAAGGAAACCACTGT	2220
Db	2574	TATGATGAGGGGATGCTTTTAAAGTGCATGTCTTGGGGCTGGGAAGGAAACCACTGT	2633
Qy	2221	AACATAGCCGAAACAGTAGCTGCCCTGCCCAACCCCTGCCATTAATGGGGGCACATGTGTG	2280
Db	2634	AACATAGCCGAAACAGTAGCTGCCCTGCCCAACCCCTGCCATTAATGGGGGCACATGTGTG	2693
Qy	2281	GTCAACGGCGAGTCTTTTACGTGCGTCTGCAAGGAAGGCTGGGAGGGGCCCACTGTGTCT	2340
Db	2694	GTCAACGGCGAGTCTTTTACGTGCGTCTGCAAGGAAGGCTGGGAGGGGCCCACTGTGTCT	2753
Qy	2341	CAGAAATCAATGACTGCGAGCCCTCATCTGTTTACAACAGCGGCACTGTGTGGATGGA	2400
Db	2754	CAGAAATCAATGACTGCGAGCCCTCATCTGTTTACAACAGCGGCACTGTGTGGATGGA	2813
Qy	2401	GACAACTGGTACCGGTGCGAATGTGCCCGGGTTTGTGGGCCCGGACTGCAGAAATAAAC	2460
Db	2814	GACAACTGGTACCGGTGCGAATGTGCCCGGGTTTGTGGGCCCGGACTGCAGAAATAAAC	2873
Qy	2461	ATCAATGAATGCCAGTCTTTCACCTTGTCCCTTTGGAGCGACCTGTGTGGATGAGATCAAT	2520
Db	2874	ATCAATGAATGCCAGTCTTTCACCTTGTCCCTTTGGAGCGACCTGTGTGGATGAGATCAAT	2933
Qy	2521	GGCTACCGGTGTCTGCCCTCAGGGCACAGTGGTGCCAAAGTCCAGGAAGTTTCAGGG	2580
Db	2934	GGCTACCGGTGTCTGCCCTCAGGGCACAGTGGTGCCAAAGTCCAGGAAGTTTCAGGG	2993
Qy	2581	AGACCTTGCAATCAACATGGGGAGTGTGATACCAAGTGGGGCCAAATGGGATGATGACTGT	2640
Db	2994	AGACCTTGCAATCAACATGGGGAGTGTGATACCAAGTGGGGCCAAATGGGATGATGACTGT	3053
Qy	2641	AATACCTGCCAGTCCCTGAAATGGAACGGATCGCTGTCTCAAAAGTGTGGTGTGGCCCTCGA	2700
Db	3054	AATACCTGCCAGTCCCTGAAATGGAACGGATCGCTGTCTCAAAAGTGTGGTGTGGCCCTCGA	3113
Qy	2701	CCTTGCTGTCTCCAAAGGGGCAAGGAGTGCCCGAGCGGGGAGAGCTGCATCCCCATC	2760
Db	3114	CCTTGCTGTCTCCAAAGGGGCAAGGAGTGCCCGAGCGGGGAGAGCTGCATCCCCATC	3173
Qy	2761	CTGAGACAGGAGTCTTCCCTCCACCTGCACCTGTGTGGCGAGTGTCCGCTCTTCAGT	2820
Db	3174	CTGAGACAGGAGTCTTCCCTCCACCTGCACCTGTGTGGCGAGTGTCCGCTCTTCAGT	3233
Qy	2821	CTCAGCGCGTGAAGACAAAGTGCACCTCTGACTCTCTATTACCAGGATAAATCTGTGCAAC	2880
Db	3234	CTCAGCGCGTGAAGACAAAGTGCACCTCTGACTCTCTATTACCAGGATAAATCTGTGCAAC	3293
Qy	2881	ATCAATTTACCTTTAAACAGGAGATGATGTACACGAGTCTTACTACGGAGCAATTTGC	2940
Db	3294	ATCAATTTACCTTTAAACAGGAGATGATGTACACGAGTCTTACTACGGAGCAATTTGC	3353
Qy	2941	AGTGAATTTAGGAATTTTGAATATTTTGAAGAAATGTTTCCGCTGAATATTCAATCTACATC	3000
Db	3354	AGTGAATTTAGGAATTTTGAATATTTTGAAGAAATGTTTCCGCTGAATATTCAATCTACATC	3413
Qy	3001	GCTTGGCAGCTTCCCTTCCAGCGAAACAAATGAATAACATGTGGCCATTTCTGTGGAAGAT	3060

Db 3414 GCTTGGAGCCTTCCCTTTCAGCGAACAATGAAATACATGTGGCCATTTCTGCTGAAGAT 3473
Qy 3061 ATACGGGATGATGGAAACCCGATCAAGGAAATCACTGACAAAATAATCGATCTTTGTTAGT 3120
Db 3474 ATACGGGATGATGGAAACCCGATCAAGGAAATCACTGACAAAATAATCGATCTTTGTTAGT 3533
Qy 3121 AAAGCTGATGAAACAGCTGCTGATTTGCTGCGCTTGCAGAGTAAGAGTTTCAGAGCGG 3180
Db 3534 AAAGCTGATGAAACAGCTGCTGATTTGCTGCGCTTGCAGAGTAAGAGTTTCAGAGCGG 3593
Qy 3181 CCTCTGAAGAACAGAAACAGATTTCTTGTTCCTTGTGAGCTCTGTCTTAACTGTGCT 3240
Db 3594 CCTCTGAAGAACAGAAACAGATTTCTTGTTCCTTGTGAGCTCTGTCTTAACTGTGCT 3653
Qy 3241 TGGATCTGTTGCTGTTGAGCGGCTTTACTGTGCTGCTGCGGAAGCGCGGAAACCGGCG 3300
Db 3654 TGGATCTGTTGCTGTTGAGCGGCTTTACTGTGCTGCTGCGGAAGCGCGGAAACCGGCG 3713
Qy 3301 AGCCACACACTCAGCTCTGAGGACAAACACCAACCAACGTCGCGGAGCAGCTGAAC 3360
Db 3714 AGCCACACACTCAGCTCTGAGGACAAACACCAACCAACGTCGCGGAGCAGCTGAAC 3773
Qy 3361 CAGATCAAAAACCCCATTTAGAAACATGGGGCCCAACACGGTCCCATCAAGGATTACGAG 3420
Db 3774 CAGATCAAAAACCCCATTTAGAAACATGGGGCCCAACACGGTCCCATCAAGGATTACGAG 3833
Qy 3421 AACAGAACTCCAAATGTCTTAAATAAGGACACACAATTTCTGAAGTAGAAGGAGCAGC 3480
Db 3834 AACAGAACTCCAAATGTCTTAAATAAGGACACACAATTTCTGAAGTAGAAGGAGCAGC 3893
Qy 3481 ATGACAAAACACGAGGAAAGCCGGTTTGGCAGCAGCGCGGCTATAGCTGTGTAGAC 3540
Db 3894 ATGACAAAACACGAGGAAAGCCGGTTTGGCAGCAGCGCGGCTATAGCTGTGTAGAC 3593
Qy 3541 AGAGAAGAGAGGCCCCCAACGCGCACCGCACCAACCAACCAACCTGACCAACCAACAG 3600
Db 3954 AGAGAAGAGAGGCCCCCAACGCGCACCGCACCAACCAACCAACCTGACCAACCAACAG 4013
Qy 3601 GACAAACAGAGACTTGGAAAGTGCCAGAGCTTAAACCGAATGGAATGATCATCGTATAG 3657
Db 4014 GACAAACAGAGACTTGGAAAGTGCCAGAGCTTAAACCGAATGGAATGATCATCGTATAG 4070

RESULT 4
AR482321
LOCUS 5590-69-1 mRNA linear PAT 14-MAY-2004
DEFINITION Sequence 1 from patent US 6,703,166
ACCESSION AR482321
VERSION AR482321.1 GI:47244642
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 5590)
AUTHORS Li, L., Hood, L., Krantz, I.D. and Spinner, N.B.
TITLE Methods of diagnosing alagille syndrome
JOURNAL Patent: US 6703198-A 1 09-MAR-2004
The University of Washington and The Children's Hospital of Philadelphia; Seattle, WA
FEATURES
source Location/Qualifiers
1. 5590
/organism="unknown"
/mol_type="mRNA"
ORIGIN
Query Match 99.98; Score 3654.2; DB 6; Length 5590;
Best Local Similarity 99.98; Pred. No. 0;
Matches 3653; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGCGTTCCCGACGACGCGCGCGCTCGGGCGCCCTAAGCCTTCCTGCTCGCCCTG 60
Db 414 ATGCGTTCCCGACGACGCGCGCGCTCGGGCGCCCTAAGCCTTCCTGCTCGCCCTG 473

Qy 61 CTCTGTGCCCCGAGCCAAAGGTGTGTGGGGCTTCGGGTCTGAGTTGGAGATCCTG 120
Db 474 CTCTGTGCCCCGAGCCAAAGGTGTGTGGGGCTTCGGGTCTGAGTTGGAGATCCTG 533
Qy 121 TCCATGACAGAACTGAAACGGGAGCTGCAGAAACGGGAACCTGTCGGCGGGCCCGGAAAC 180
Db 534 TCCATGACAGAACTGAAACGGGAGCTGCAGAAACGGGAACCTGTCGGCGGGCCCGGAAAC 593
Qy 181 CCGGGAGACCGCAAGTGCACCCCGCGAGTGTGACACATCTTCAAAGTGTGCTCAAG 240
Db 594 CCGGGAGACCGCAAGTGCACCCCGCGAGTGTGACACATCTTCAAAGTGTGCTCAAG 653
Qy 241 GAGTATCAGTCCCGCTGACCGCGGGGGCCCTGACGCTTCGCTCAGGGTTCACGCT 300
Db 654 GAGTATCAGTCCCGCTGACCGCGGGGGCCCTGACGCTTCGCTCAGGGTTCACGCT 713
Qy 301 GTCATCGGGGGCAACACTTCAACCTCAAGCCGAGCGCGCAACGACCGCAACCGCATC 360
Db 714 GTCATCGGGGGCAACACTTCAACCTCAAGCCGAGCGCGCAACGACCGCAACCGCATC 773
Qy 361 GTCTGTCTTTTCAAGTTTCGCTTCGCGAGGTCTCTATACGTTGTGTGGAGGGGTGGAT 420
Db 774 GTCTGTCTTTTCAAGTTTCGCTTCGCGAGGTCTCTATACGTTGTGTGGAGGGGTGGAT 833
Qy 421 TCCAGTAAATGACACCGCTTCAACCTGACAGTATTTATGAAAAGGCTTCTCACTCGGGCATG 480
Db 834 TCCAGTAAATGACACCGCTTCAACCTGACAGTATTTATGAAAAGGCTTCTCACTCGGGCATG 893
Qy 481 ATCAACCCCGAGCGGAGTGGCAGAGCGCTGAAGCAGAACACGGCGCTTGCCTTTGAG 540
Db 894 ATCAACCCCGAGCGGAGTGGCAGAGCGCTGAAGCAGAACACGGCGCTTGCCTTTGAG 953
Qy 541 TATCAGATCCGCTGACCTGTGTGATGACTACTATATGGCTTTGGCTGTAAATAGTTCTGC 600
Db 954 TATCAGATCCGCTGACCTGTGTGATGACTACTATATGGCTTTGGCTGTAAATAGTTCTGC 1013
Qy 601 CGCCCGAGAGATGACTTTCTTTGACACTATGCTGTGACAGAGTGGCAACAAACCTTGC 660
Db 1014 CGCCCGAGAGATGACTTTCTTTGACACTATGCTGTGACAGAGTGGCAACAAACCTTGC 1073
Qy 661 ATGGAAGGCTGGATGGGCGGCGGCTTAAACAGAGCTATTTGCGCAAGGCTCGAGTCT 720
Db 1074 ATGGAAGGCTGGATGGGCGGCGGCTTAAACAGAGCTATTTGCGCAAGGCTCGAGTCT 1133
Qy 721 AAGCATGGGTCTTGCAGAACTCCAGAGTGCAGTGCAGGTGCGAGTAYGGCTGGCAAGGCTG 780
Db 1134 AAGCATGGGTCTTGCAGAACTCCAGAGTGCAGTGCAGGTGCGAGTAYGGCTGGCAAGGCTG 1193
Qy 781 TACTGTGATAAGTGCATCCACACCGGGGATGCTCCAGCGCATCTGTAAATGAGCCCTCG 840
Db 1194 TACTGTGATAAGTGCATCCACACCGGGGATGCTCCAGCGCATCTGTAAATGAGCCCTCG 1253
Qy 841 CAGTGTCTGTGAGACCAACTCGGGCGGCGCAGCTCTGTGACAAAGATCTCAATTAAGT 900
Db 1254 CAGTGTCTGTGAGACCAACTCGGGCGGCGCAGCTCTGTGACAAAGATCTCAATTAAGT 1313
Qy 901 GGGACTCATCAGCCGCTGTCTCAACGGGGGAACTTTGTAGCAACACAGCCCTCGCAATAT 960
Db 1314 GGGACTCATCAGCCGCTGTCTCAACGGGGGAACTTTGTAGCAACACAGCCCTCGCAATAT 1373
Qy 961 CAGTGTCTGTGAGGAGTATTCAGGACCCCACTGTGAAATTTGTGAGAGCGCCCTGC 1020
Db 1374 CAGTGTCTGTGAGGAGTATTCAGGACCCCACTGTGAAATTTGTGAGAGCGCCCTGC 1433
Qy 1021 CTCTCTGATCCCTGTGACACAGAGCGCTGTAAAGGAGACCTCCCTGGGCTTTGAGTGT 1080
Db 1434 CTCTCTGATCCCTGTGACACAGAGCGCTGTAAAGGAGACCTCCCTGGGCTTTGAGTGT 1493
Qy 1081 GAGTGTTCCTCCAGGCTGGACCGGCGGCGCAGCTCTTACAAAACATGATGATGCTGTCTCT 1140
Db 1494 GAGTGTTCCTCCAGGCTGGACCGGCGGCGCAGCTCTTACAAAACATGATGATGCTGTCTCT 1553
Qy 1141 AATAACGTGTTCCCAACGGGGGCGACCTGTCAGGACCTGTTTAAACGGATTTAAGTGTGTGC 1200

1554	Db	 AATAA	 CTGTTCC	 ACGCGGG	 CACCTG	 CCAGGAC	 CTGGTTA	 ACGGA	 TTTAAGT	 GTGTGC	1613
1201	Qy	 CCCCA	 CAGTGG	 ACGTGG	 AAGACGT	 GCAGT	 TAGATG	 CAATG	 AATGTG	 AGGCCA	1260
1614	Db	 CCCCA	 CAGTGG	 ACGTGG	 AAGACGT	 GCAGT	 TAGATG	 CAATG	 AATGTG	 AGGCCA	1673
1261	Qy	 TGTTAA	 ACCCAAA	 TCTCTG	 TAAAGAA	 TCTCAIT	 TGCCAGT	 TACTACT	 GCGACT	 GTTTCCC	1320
1674	Db	 TGTTAA	 ACCCAAA	 TCTCTG	 TAAAGAA	 TCTCAIT	 TGCCAGT	 TACTACT	 GCGACT	 GTTTCCC	1733
1321	Qy	 GGCTGG	 AATGGG	 TCAGAA	 TTGTGCA	 TATAATA	 TATGA	 CTGCC	 TTGGC	 CAGTGTCA	1380
1734	Db	 GGCTGG	 AATGGG	 TCAGAA	 TTGTGCA	 TATAATA	 TATGA	 CTGCC	 TTGGC	 CAGTGTCA	1793
1381	Qy	 GAGCCT	 CTCTGG	 GATTTGG	 TTAATGG	 TTATCG	 CTGTAT	 CTGTG	 CACCTG	 GCGCTATG	1440
1794	Db	 GAGCCT	 CTCTGG	 GATTTGG	 TTAATGG	 TTATCG	 CTGTAT	 CTGTG	 CACCTG	 GCGCTATG	1853
1441	Qy	 GGCGAT	 CACTGT	 GAGAG	 CATCGAT	 GAATG	 GCCAG	 CAACCC	 CTGTTGA	 ATGGGGT	1500
1854	Db	 GGCGAT	 CACTGT	 GAGAG	 CATCGAT	 GAATG	 GCCAG	 CAACCC	 CTGTTGA	 ATGGGGT	1913
1501	Qy	 CACTGT	 CAGAA	 TCAAC	 GATTC	 CCAGT	 GTCTGT	 CCCCA	 TGTTTCT	 CTGAAAC	1560
1914	Db	 CACTGT	 CAGAA	 TCAAC	 GATTC	 CCAGT	 GTCTGT	 CCCCA	 TGTTTCT	 CTGAAAC	1973
1561	Qy	 CTCTGT	 CAGTGG	 CATCG	 ATTTATG	 TGACCT	 TAATCC	 CTGCC	 CAGAC	 CGTGCC	1620
1974	Db	 CTCTGT	 CAGTGG	 CATCG	 ATTTATG	 TGACCT	 TAATCC	 CTGCC	 CAGAC	 CGTGCC	2033
1621	Qy	 TACAA	 CCGTGC	 CAGTAC	 TATTTCT	 GCAG	 TGCCCC	 GAGCA	 CTATG	 AGGCCA	1680
2034	Db	 TACAA	 CCGTGC	 CAGTAC	 TATTTCT	 GCAG	 TGCCCC	 GAGCA	 CTATG	 AGGCCA	2093
1681	Qy	 TCA	 CACTG	 AAAGCA	 CCTCG	 CA	 CCCCCT	 GTGA	 GTGAT	 TGACAG	1740
2094	Db	 TCA	 CACTG	 AAAGCA	 CCTCG	 CA	 CCCCCT	 GTGA	 GTGAT	 TGACAG	2153
1741	Qy	 GCCATG	 CTTCCA	 ACGAC	 ACACCTG	 AAGGG	 TGGGTAT	 TTTCTC	 TCCAC	 CGTGTGGT	1800
2154	Db	 GCCATG	 CTTCCA	 ACGAC	 ACACCTG	 AAGGG	 TGGGTAT	 TTTCTC	 TCCAC	 CGTGTGGT	2213
1801	Qy	 CCTC	 ACGGG	 AAGTCA	 AGAGT	 TCAGT	 CGGGG	 CAAA	 TTCACT	 GTGTACT	1860
2214	Db	 CCTC	 ACGGG	 AAGTCA	 AGAGT	 TCAGT	 CGGGG	 CAAA	 TTCACT	 GTGTACT	2273
1861	Qy	 TTCA	 CGGGAA	 CATACT	 GCCAT	 GAAAA	 TATTA	 TGA	 CTGTG	 AGCAAC	1920
2274	Db	 TTCA	 CGGGAA	 CATACT	 GCCAT	 GAAAA	 TATTA	 TGA	 CTGTG	 AGCAAC	2333
1921	Qy	 GGTGG	 CACTG	 CATG	 GTGTCA	 ACTCCT	 CAAGT	 GCACT	 GTGTAG	 CGCTGG	1980
2334	Db	 GGTGG	 CACTG	 CATG	 GTGTCA	 ACTCCT	 CAAGT	 GCACT	 GTGTAG	 CGCTGG	2393
1981	Qy	 GGGGCT	 TACTGT	 GAAA	 CAAA	 TATTA	 ATG	 ACTG	 CAGC	 CAACCC	2040

2634	AA	CATAGCCGGAACAGTAGCTGCCCTGCCCAACCCCTGCCATAATGGGGGCACATGTGCTG	2699
2281	GT	CAAGCGGCGAGTCTTTTACGTGCGCTGTGCAAGGAAGGCTGGAGGGGGCCATCTGTGCT	2340
2694	GT	CAAGCGGCGAGTCTTTACGTGCGCTGTGCAAGGAAGGCTGGAGGGGGCCATCTGTGCT	2753
2341	CAGA	ATACCAATGACTGCGAGCCCTCATCTCTGTTAACAACGGGCACCTGTGTGGATGGA	2400
2754	CAGA	ATACCAATGACTGCGAGCCCTCATCTCTGTTAACAACGGGCACCTGTGTGGATGGA	2813
2401	GACA	CTGTGTACCGGTGGGAATGTGCCCGGGTTTTGCTGGGCCCGACTGCAGAAATAAC	2460
2814	GACA	CTGTGTACCGGTGGGAATGTGCCCGGGTTTTGCTGGGCCCGACTGCAGAAATAAC	2873
2461	ATCA	ATGAATGCCAGTCTTCACCTTGTGCTTTGGAGCGACCTGTGTGGATGAGATCAAT	2520
2874	ATCA	ATGAATGCCAGTCTTCACCTTGTGCTTTGGAGCGACCTGTGTGGATGAGATCAAT	2933
2521	GGCT	ACGGTGTGTCTGCCCTCCAGGCGACAGTGTGTGCCAAGTGCACAGGAAGTTTCAGGG	2588
2934	GGCT	ACGGTGTGTCTGCCCTCCAGGCGCGACAGTGTGTGCCAAGTGCACAGGAAGTTTCAGGG	2993
2581	AGAC	CTTGCAATCAACATGGGAGTGTGATACAGATGGGGGCCAAATGGGATGATGACTGT	2640
2994	AGAC	CTTGCAATCAACATGGGAGTGTGATACAGATGGGGGCCAAATGGGATGATGACTGT	3053
2641	AATAC	TGCGAGTGCCTTGAATGCAACGGATCGCCTGTCTCAAAAGTCTGTGTGGCCCTCGA	2700
3054	AATAC	TGCGAGTGCCTTGAATGCAACGGATCGCCTGTCTCAAAAGTCTGTGTGGCCCTCGA	3113
2701	CTTG	CGCTGTCTCAAAAGGGCACAGCGAGTGCCTCCAGCGGGCAGAGCTGCATCCCCATC	2760
3114	CTTG	CGCTGTCTCAAAAGGGCACAGCGAGTGCCTCCAGCGGGCAGAGCTGCATCCCCATC	3173
2761	CTGA	CGACAGTGTCTGCTCAACCCCTGCACCTGTGTGGCGGAGTGTCCGCTCTTCAGT	2820
3174	CTGA	CGACAGTGTCTGCTCAACCCCTGCACCTGTGTGGCGGAGTGTCCGCTCTTCAGT	3233
2821	CTCAG	CGCGTGAAGACAAAGTCACCTCTGACTCTTATTAACAGGATAACTGTGCGAATC	2880
3234	CTCAG	CGCGTGAAGACAAAGTCACCTCTGACTCTTATTAACAGGATAACTGTGCGAATC	3293
2881	ATCA	CATTTACCTTTAAACAGGAGATGATCTCACAGGTCTTACTACGGAGCACATTTGC	2940
3294	ATCA	CATTTACCTTTAAACAGGAGATGATCTCACAGGTCTTACTACGGAGCACATTTGC	3353
2941	AGTG	AATGAGGAATTTGGAATATTTTGAAGAAATGTTTCCGCTGNAATATTAATCTCAATC	3000
3354	AGTG	AATGAGGAATTTGGAATATTTTGAAGAAATGTTTCCGCTGNAATATTAATCTCAATC	3413
3001	GCTT	GCGAGCTTCCCCTTCAGCGAAACAAATGAATACATGTGGGCCATTTCTGCTGAAGAT	3060
3414	GCTT	GCGAGCTTCCCCTTCAGCGAAACAAATGAATACATGTGGGCCATTTCTGCTGAAGAT	3473
3061	ATAC	GCGGATGATGGGAAACCGGATCAAGGAAATCACTGAACAAAATAATCGATCTTTGTAGT	3120
3474	ATAC	GCGGATGATGGGAAACCGGATCAAGGAAATCACTGAACAAAATAATCGATCTTTGTAGT	3533
3121	AAAC	GATGGAATAACAGCTCGCTGATTTGCTGCGCTTCAGAAATGAAGAGTTTCAGAGCGG	3180
3534	AAAC	GATGGAATAACAGCTCGCTGATTTGCTGCGCTTCAGAAATGAAGAGTTTCAGAGCGG	3593
3181	CCTC	TGAAGAACAGAACAGATTTTCTTTGTTCCCTTGTCTGAGCTCTGTCTTAACTGTGGCT	3240
3594	CCTC	TGAAGAACAGAACAGATTTTCTTTGTTCCCTTGTCTGAGCTCTGTCTTAACTGTGGCT	3653
3241	TGGA	TCTGTGTCAGCGCTTCTACTGTGTGCTTCGCGAAGCGCGCGGAAGCCGGGC	3300
3654	TGGA	TCTGTGTCAGCGCTTCTACTGTGTGCTTCGCGAAGCGCGCGGAAGCCGGGC	3713
3301	AGCC	CACACACTCAGCTCTGAGGACAAACACCAACCAACAGTCGTGCGGAGCAGCTGAAC	3360
3714	AGCC	CACACACTCAGCTCTGAGGACAAACCAACCAACAGTCGTGCGGAGCAGCTGAAC	3773

Db	1074		ATGGAAGGCTGGATGGGCCCGGAATGTAAACAGAGCTATTTTGCCGACAAGGCTCGAGTCT	1133
Qy	721		AAGATGGGCTTTGCAAACTCCAGGTGATCTGCAGGTGCCAGTAYGGCTGGCAAGGCTTG	780
Db	1134		AAGCATGGGCTTTGCAAACTCCAGGTGATCTGCAGGTGCCAGTATGGCTGGCAAGGCTTG	1193
Qy	781		TACTGTGATAAGTGCATCCACACCCCGGATGGCTCCAGCGCATCTGTAATGAGCCCTGG	840
Db	1194		TACTGTGATAAGTGCATCCACACCCCGGATGGCTCCAGCGCATCTGTAATGAGCCCTGG	1253
Qy	841		CAGTGCCCTCTGTGAGACCAACTGGGGCGGCGAGCTCTGTGACAAAAGATCTCAATTA	900
Db	1254		CAGTGCCCTCTGTGAGACCAACTGGGGCGGCGAGCTCTGTGACAAAAGATCTCAATTA	1313
Qy	901		GGGACTCATCAGCGGTGTCTCAACGGGGGAACTTGTAGCAACACAGGCGCTGCAAAATAT	960
Db	1314		GGGACTCATCAGCGGTGTCTCAACGGGGGAACTTGTAGCAACACAGGCGCTGCAAAATAT	1373
Qy	961		CAGTGTTCCTGCCCTGAGGGGTATTACAGGACCCAACTGTGAAATTTGCTGAGCAAGCTGC	1020
Db	1374		CAGTGTTCCTGCCCTGAGGGGTATTACAGGACCCAACTGTGAAATTTGCTGAGCAAGCTGC	1433
Qy	1021		CTCTCTGATCCCTGTCAACACAGAGGCAAGTGTAAAGAGACCTCCCTGGGCTTTGAGTGT	1080
Db	1434		CTCTCTGATCCCTGTCAACACAGAGGCAAGTGTAAAGAGACCTCCCTGGGCTTTGAGTGT	1493
Qy	1081		GAGTGTTCCTGCCCTGAGGGGTATTACAGGACCCAACTGTGAAATTTGCTGAGCAAGCTGC	1140
Db	1494		GAGTGTTCCTGCCCTGAGGGGTATTACAGGACCCAACTGTGAAATTTGCTGAGCAAGCTGC	1553
Qy	1141		AATAAATCTTCCACAGGGGCACTGCCAGGACCTGGTTAAAGGATTTAAAGTGTGTGTC	1200
Db	1554		AATAAATCTTCCACAGGGGCACTGCCAGGACCTGGTTAAAGGATTTAAAGTGTGTGTC	1613
Qy	1201		CCCCCAAGTGGGACTGGGAAAACTGGCCAGTGTAGATGCAAAATGAATGTAGGGCCAAACCT	1260
Db	1614		CCCCCAAGTGGGACTGGGAAAACTGGCCAGTGTAGATGCAAAATGAATGTAGGGCCAAACCT	1673
Qy	1261		TGTGTAAACGGCAAACTCTGTAGAAATCTATTCGCGAGCTACTCTCGGAGTGTCTTCCC	1320
Db	1674		TGTGTAAACGGCAAACTCTGTAGAAATCTATTCGCGAGCTACTCTCGGAGTGTCTTCCC	1733
Qy	1321		GGCTGGATGGGTGAGAAATTTGACATAAAATTAATGACTGCCTTGGCCAGTGTGAGAAAT	1380
Db	1734		GGCTGGATGGGTGAGAAATTTGACATAAAATTAATGACTGCCTTGGCCAGTGTGAGAAAT	1793
Qy	1381		GAGCCCTCCTGTGGGGAATTTGGTTAATCGGTATCGCTGTATCTGTCCACCTGGGCTATGCA	1440
Db	1794		GAGCCCTCCTGTGGGGAATTTGGTTAATCGGTATCGCTGTATCTGTCCACCTGGGCTATGCA	1853
Qy	1441		GGCGATCACTGTGAGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTGAATGGGGGT	1500
Db	1854		GGCGATCACTGTGAGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTGAATGGGGGT	1913
Qy	1501		CACGTGAGAAATGAAATCAACAGATTCAGGTGCTGTGTCCCACTGTCTTCTCGGAAAC	1560
Db	1914		CACGTGAGAAATGAAATCAACAGATTCAGGTGCTGTGTCCCACTGTCTTCTCGGAAAC	1973
Qy	1561		CTCTGTCACTGTGACATCGATTAATGTGAGCCTAATCCCTGCCAGAACCGTGGCCAGTGC	1620
Db	1974		CTCTGTCACTGTGACATCGATTAATGTGAGCCTAATCCCTGCCAGAACCGTGGCCAGTGC	2033
Qy	1621		TACAAACCGTGCAGTGACTATTTCTGCAAGTGGCCCGAGGACTATGAGGGCAAGACTGC	1680
Db	2034		TACAAACCGTGCAGTGACTATTTCTGCAAGTGGCCCGAGGACTATGAGGGCAAGACTGC	2093
Qy	1681		TCACACCTGAAAGACCACTGGCGGACGACCCCTGTGAAAGTATTGACAGCTGCACAGTG	1740
Db	2094		TCACACCTGAAAGACCACTGGCGGACGACCCCTGTGAAAGTATTGACAGCTGCACAGTG	2153
Qy	1741		GCCATGGCTTCCAAAGCACACCTGAAGGGGTGGGTATATTTCTCCACAGCTGTGGT	1800

Db	2154		GCCATGGCTTCCAAAGCACACACTGAAGGGGTGGGTATATTTCTCCACAGCTGTGTGGT	2213
Qy	1801		CCTTCACGGGAAGTGCAGAGTCACTCGGAGGCAAAATTCACCTGTGATCTGTAAACAAAGGC	1860
Db	2214		CCTTCACGGGAAGTGCAGAGTCACTCGGAGGCAAAATTCACCTGTGATCTGTAAACAAAGGC	2273
Qy	1861		TTTCAACGGGAACATATCTGCCATGAAAAATATTAATGACTGTGAGAGCAACCCCTTTAGAAAC	1920
Db	2274		TTTCAACGGGAACATATCTGCCATGAAAAATATTAATGACTGTGAGAGCAACCCCTTTAGAAAC	2333
Qy	1921		GGTGGCACTTGATTCGATCGATGGTCACTCTCAAGTGCATCTGTAGTGAAGCTGGGAG	1980
Db	2334		GGTGGCACTTGATTCGATCGATGGTCACTCTCAAGTGCATCTGTAGTGAAGCTGGGAG	2393
Qy	1981		GGGGCCCTACTGTGAAACCAATATTAATGACTGTGAGAGCAACCCCTGCCACAAATGGGGC	2040
Db	2394		GGGGCCCTACTGTGAAACCAATATTAATGACTGTGAGAGCAACCCCTGCCACAAATGGGGC	2453
Qy	2041		ACGTGTGCGGACCTGGTCAATGACTTCTACTGTGACTGTAAAAATGGGTGGAAAGAAAG	2100
Db	2454		ACGTGTGCGGACCTGGTCAATGACTTCTACTGTGACTGTAAAAATGGGTGGAAAGAAAG	2513
Qy	2101		ACCTGCCACTCAGTGCAGTCACTGATGAGGCAACCGTGCACACACGCTGCACTGC	2160
Db	2514		ACCTGCCACTCAGTGCAGTCACTGATGAGGCAACCGTGCACACACGCTGCACTGC	2573
Qy	2161		TATGATGAGGGGATGCTTTAAAGTGCATGTGTCTCTGGCGCTGGGAAAGAAACCACTGT	2220
Db	2574		TATGATGAGGGGATGCTTTAAAGTGCATGTGTCTCTGGCGCTGGGAAAGAAACCACTGT	2633
Qy	2221		AACATAGCCCCGAAACAGTAGTCCCTGCCAACCCCTGCCCATATGGGGGCACATGTGTG	2280
Db	2634		AACATAGCCCCGAAACAGTAGTCCCTGCCAACCCCTGCCCATATGGGGGCACATGTGTG	2693
Qy	2281		GTCAAACGGGAGTCTTTTACGTGCGTCTGCAAGGAGGCTGGGAGGGGCCCCATCTGTGCT	2340
Db	2694		GTCAAACGGGAGTCTTTTACGTGCGTCTGCAAGGAGGCTGGGAGGGGCCCCATCTGTGCT	2753
Qy	2341		CAGAAATACCAATGACTGACGCTCATCCCTGTTTACAAACAGCGGACCTGTGTGGATGGA	2400
Db	2754		CAGAAATACCAATGACTGACGCTCATCCCTGTTTACAAACAGCGGACCTGTGTGGATGGA	2813
Qy	2401		GACAACTGTGTACCGGTGCGAATGTGCCCGGGTTTTGCTGGGCCCGCATCAGAAATAAAC	2460
Db	2814		GACAACTGTGTACCGGTGCGAATGTGCCCGGGTTTTGCTGGGCCCGCATCAGAAATAAAC	2873
Qy	2461		ATCAATGAATGCCAGTCTTCACTTGTGCGCTTTGGAGCGACCTGTGTGGATGAGATCAAT	2520
Db	2874		ATCAATGAATGCCAGTCTTCACTTGTGCGCTTTGGAGCGACCTGTGTGGATGAGATCAAT	2933
Qy	2521		GGCTACCGGTGTGTCTGCCCTCCAGGGCACAGTGGTGCCTCAAGTGCACAGGAAATTCAGGG	2580
Db	2934		GGCTACCGGTGTGTCTGCCCTCCAGGGCACAGTGGTGCCTCAAGTGCACAGGAAATTCAGGG	2993
Qy	2581		AGACCTTTGCATCACTATGGGGAGTGTGATACCAAGTGGGGCCAAATGGGATGATGACTGT	2640
Db	2994		AGACCTTTGCATCACTATGGGGAGTGTGATACCAAGTGGGGCCAAATGGGATGATGACTGT	3053
Qy	2641		AATACCTGCCAGTGCCTGAATGAGCGGATCGCTGTCTCAAGTGTGGTGGGCCCTCGA	2700
Db	3054		AATACCTGCCAGTGCCTGAATGAGCGGATCGCTGTCTCAAGTGTGGTGGGCCCTCGA	3113
Qy	2701		CCTTGGCTGTCCCAAAAGGGGCAAGCGAGTGGCCCGAGGGGAGAGCTGCATCCCCATC	2760
Db	3114		CCTTGGCTGTCCCAAAAGGGGCAAGCGAGTGGCCCGAGGGGAGAGCTGCATCCCCATC	3173
Qy	2761		CTGGAACGACAGTGTCTTCTGCTCCACCCCTGCACTGGTGTGGGCGAGTGTGGTCTTCAGT	2820
Db	3174		CTGGAACGACAGTGTCTTCTGCTCCACCCCTGCACTGGTGTGGGCGAGTGTGGTCTTCAGT	3233
Qy	2821		CTCCAGCGGTGAAGACAAAGTGCACCTCTGACTCTCTATTACCAAGGATTAATGTGGGAAAC	2880
Db	3234		CTCCAGCGGTGAAGACAAAGTGCACCTCTGACTCTCTATTACCAAGGATTAATGTGGGAAAC	3293

QY 2881 ATGACATTTACCTTTAAACAGGAGATGATGTACCAGGTCTTACTACGGAGCAATTTGC 2940
DB 3294 ATGACATTTACCTTTAAACAGGAGATGATGTACCAGGTCTTACTACGGAGCAATTTGC 3353
QY 2941 AGTGAAATTGAGGAATTTGAATATTTTGAAGAAATGTTTCCGCTGAATATTAATCAATCATC 3000
DB 3354 AGTGAAATTGAGGAATTTGAATATTTTGAAGAAATGTTTCCGCTGAATATTAATCAATCATC 3413
QY 3001 GCTTGCAGGCTTCCCTTTCAGCGAAACAAATGAAATACATGTGGCCATTTCTGTCTGAAGAT 3060
DB 3414 GCTTGCAGGCTTCCCTTTCAGCGAAACAAATGAAATACATGTGGCCATTTCTGTCTGAAGAT 3473
QY 3061 ATACGGATGATGGGAACCGGATCAAGGAATCACTGACAAAATAATTCGATCTTGTAGT 3120
DB 3474 ATACGGATGATGGGAACCGGATCAAGGAATCACTGACAAAATAATTCGATCTTGTAGT 3533
QY 3121 AAACGATGATGAAACAGCTCGCTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180
DB 3534 AAACGATGATGAAACAGCTCGCTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3593
QY 3181 CCTCTGAAGAACAGCAAGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3240
DB 3594 CCTCTGAAGAACAGCAAGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3653
QY 3241 TGGATCTGTTGCTTGTGTGAGCGGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3300
DB 3654 TGGATCTGTTGCTTGTGTGAGCGGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3713
QY 3301 AGCCACACACACTCAGGCTCTGAGGACAAACACCAACCAACCAACCAACCAACCAACCAACCA 3360
DB 3714 AGCCACACACACTCAGGCTCTGAGGACAAACACCAACCAACCAACCAACCAACCAACCAACCA 3773
QY 3361 CAGATCAAAAACCCCATTTGAGAAACATGGGGCCCAACACCGTCCCATCAAGGATTAACGAG 3420
DB 3774 CAGATCAAAAACCCCATTTGAGAAACATGGGGCCCAACACCGTCCCATCAAGGATTAACGAG 3833
QY 3421 AACAGAACTCCAAATGCTTAAATAAGGACACACAATTTCTGAAGTAGAAGAGGAGCAC 3480
DB 3834 AACAGAACTCCAAATGCTTAAATAAGGACACACAATTTCTGAAGTAGAAGAGGAGCAC 3893
QY 3481 ATGACAAAACACACGAGGAAGCCCGTTTGGCAAGCAGCGCGGTATACGCTGGTAGAC 3540
DB 3894 ATGACAAAACACACGAGGAAGCCCGTTTGGCAAGCAGCGCGGTATACGCTGGTAGAC 3953
QY 3541 AGAGAAGAGAGGCCCCCAACGCGCACCGCAACAAAACACCAAACTGGACAAAACAAACAG 3600
DB 3954 AGAGAAGAGAGGCCCCCAACGCGCACCGCAACAAAACACCAAACTGGACAAAACAAACAG 4013
QY 3601 GACACAGAGACTTGGAAAGTGGCCAGAGCTTAAACCGAATGAGTACATCGTATAG 3657
DB 4014 GACACAGAGACTTGGAAAGTGGCCAGAGCTTAAACCGAATGAGTACATCGTATAG 4070

RESULT 6
CQ873974
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CQ873974
Sequence 393 from Patent WO2004076622.
CQ873974
CQ873974.1 GI:52747566

Homo sapiens (human)
Homo sapiens
Bukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
National Institute of Advanced Industrial Science and Technology
(JP)

Location/Qualifiers

source 1. 5896
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.9%; Score 3654.2; DB 6; Length 5896;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3653; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGCTTTCCACGACGACGCGCGCTCCGGGCGCCCTTAAGCCCTCTGCTCGCCCTG 60
DB 414 ATGCTTTCCACGACGACGCGCGCTCCGGGCGCCCTTAAGCCCTCTGCTCGCCCTG 473
QY 61 CTCTGTCTCCGCGACGACGAGGTGTGTGGGCTCGGGTCAGTTCGAGTTGGAGATCCTG 120
DB 474 CTCTGTCTCCGCGACGACGAGGTGTGTGGGCTCGGGTCAGTTCGAGTTGGAGATCCTG 533
QY 121 TCCATGACGAAACGTGAACGGGAGCTGCAGAACGGGACTGCTCGGGCGGCGCCGGAAC 180
DB 534 TCCATGACGAAACGTGAACGGGAGCTGCAGAACGGGACTGCTCGGGCGGCGCCGGAAC 593
QY 181 CCGGGAGACCGCAAGTGCACCCCGGACGAGTGTGACACATCTTCAAAGTGTCTCAAG 240
DB 594 CCGGGAGACCGCAAGTGTGACCCCGGACGAGTGTGACACATCTTCAAAGTGTCTCAAG 653
QY 241 GAGTATCAGTCCGCGCTCACGGCGGGGGCCCTGCAGCTTCCGCTCAGGGTCCACGCCT 300
DB 654 GAGTATCAGTCCGCGCTCACGGCGGGGGCCCTGCAGCTTCCGCTCAGGGTCCACGCCT 713
QY 301 GTCATCGGGGCAACACCTTCAAACCTCAAGCCAGCGCGGCAACGACCGGCAACCGCATC 360
DB 714 GTCATCGGGGCAACACCTTCAAACCTCAAGCCAGCGCGGCAACGACCGGCAACCGCATC 773
QY 361 GTCTGCTCTTCAAGTTTCCGCTCGCGAGTCTCTATACGTTGCTTGGAGGGCTGGAT 420
DB 774 GTCTGCTCTTCAAGTTTCCGCTCGCGAGTCTCTATACGTTGCTTGGAGGGCTGGAT 833
QY 421 TCCAGTAATGACACCGCTTCAACCTGACAGTATTTATTTGAAAGGCTTCTCACTCGGGCATG 480
DB 834 TCCAGTAATGACACCGCTTCAACCTGACAGTATTTATTTGAAAGGCTTCTCACTCGGGCATG 893
QY 481 ATCAACCCAGCGCGGAGTGGGAGAGCGCTGAAGCAGAAACACGGCGGTTCGCACTTTGAG 540
DB 894 ATCAACCCAGCGCGGAGTGGGAGAGCGCTGAAGCAGAAACACGGCGGTTCGCACTTTGAG 953
QY 541 TATCAGATCGCTGAGCTGTGATGACTACTATGCTGCTTGGCTTGGCTGCTGCTGCTGCTGCTG 600
DB 954 TATCAGATCGCTGAGCTGTGATGACTACTATGCTGCTTGGCTTGGCTGCTGCTGCTGCTGCTG 1013
QY 601 CGCCCCAGAGATGACTTCTTTTGGACACTATGCTGCTGACCAAGTGGCAACAAACCTTGC 660
DB 1014 CGCCCCAGAGATGACTTCTTTTGGACACTATGCTGCTGACCAAGTGGCAACAAACCTTGC 1073
QY 661 ATGGAAGGCTGATGGGCCCCGGAATGTAAACAGAGCTATTTTGGCAACAGGCTGAGTCTCT 720
DB 1074 ATGGAAGGCTGATGGGCCCCGGAATGTAAACAGAGCTATTTTGGCAACAGGCTGAGTCTCT 1133
QY 721 AAGCATGGCTCTTGAACCTCCAGGAGTGCAGGTCAGTAYGGCTGGCAAGGCTG 780
DB 1134 AAGCATGGCTCTTGAACCTCCAGGAGTGCAGGTCAGTAYGGCTGGCAAGGCTG 1193
QY 781 TACTGTGATAAGTGCATCCACACCGGGATGGCTCCAGCGCATCTGTATAGAGCCCTGG 840
DB 1194 TACTGTGATAAGTGCATCCACACCGGGATGGCTCCAGCGCATCTGTATAGAGCCCTGG 1253
QY 841 CAGTGCCTCTGTGAGACCAACTGGGCGGCGAGCTCTGTGACAAAAGATCTCAATTAATGT 900
DB 1254 CAGTGCCTCTGTGAGACCAACTGGGCGGCGAGCTCTGTGACAAAAGATCTCAATTAATGT 1313
QY 901 GGGACTCATCAGCGCTGTCTCAACGGGGGAACTTTGTAGCAACACAGGCGCTGCAATAT 960
DB 1314 GGGACTCATCAGCGCTGTCTCAACGGGGGAACTTTGTAGCAACACAGGCGCTGCAATAT 1373

QY	961	CAGTGTTCCTGCGCTGAGGGGTATTTCAGAGCCCAACTGTGAAATTCGTGACGACGCTGC	1020
DB	1374	CAGTGTTCCTGCGCTGAGGGGTATTTCAGAGCCCAACTGTGAAATTCGTGACGACGCTGC	1433
QY	1021	CTCTCTGATCCCTGTCTACAAACAGAGGCAGCTGTAAAGGAGACCTCCCTGGGCTTTGAGTGT	1080
DB	1434	CTCTCTGATCCCTGTCTACAAACAGAGGCAGCTGTAAAGGAGACCTCCCTGGGCTTTGAGTGT	1493
QY	1081	GAGTGTTCCTGAGCTGGACCGGGCCCAATGTCTTAACAAATTTGATGACTGTCTTCCT	1140
DB	1494	GAGTGTTCCTGAGCTGGACCGGGCCCAATGTCTTAACAAATTTGATGACTGTCTTCCT	1553
QY	1141	AATAACGTTCCTGAGCTGGACCGGGCCCAATGTCTTAACAAATTTGATGACTGTCTTCCT	1200
DB	1554	AATAACGTTCCTGAGCTGGACCGGGCCCAATGTCTTAACAAATTTGATGACTGTCTTCCT	1613
QY	1201	CCCCCAGTGGACTGGGAAAACTGCGAGTTAGATGCAAAATGAATGTGAGGGCCAAACCT	1260
DB	1614	CCCCCAGTGGACTGGGAAAACTGCGAGTTAGATGCAAAATGAATGTGAGGGCCAAACCT	1673
QY	1261	TGTGTAAACCCAAATCCTGTAAAGAAATCTCATTGCCAGCTACTGCGACTGTCTTCCC	1320
DB	1674	TGTGTAAACCCAAATCCTGTAAAGAAATCTCATTGCCAGCTACTGCGACTGTCTTCCC	1733
QY	1321	GGCTGGATGGGTGAGAAATTTGACATAAATATTAATGACTGCTTGGCCAGTGTCAAGAT	1380
DB	1734	GGCTGGATGGGTGAGAAATTTGACATAAATATTAATGACTGCTTGGCCAGTGTCAAGAT	1793
QY	1381	GAGCCCTCCTGTGCGGATTTGGTTAAATGGTTATCGCTGTATCTGTCCACCTGGCTATGCA	1440
DB	1794	GAGCCCTCCTGTGCGGATTTGGTTAAATGGTTATCGCTGTATCTGTCCACCTGGCTATGCA	1853
QY	1441	GGCGATCACTGTGAGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGGT	1500
DB	1854	GGCGATCACTGTGAGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGGT	1913
QY	1501	CAGTGTGAGAAATCAACAGATTCAGTGTCTGTCTGCCACTGGTTTCTCTGGAAAC	1560
DB	1914	CAGTGTGAGAAATCAACAGATTCAGTGTCTGTCTGCCACTGGTTTCTCTGGAAAC	1973
QY	1561	CTCTGTGAGTGGACATCGATTTATGTGAGCCTTAATCCCTGCCAGAACGGTGCCTCAGTGC	1620
DB	1974	CTCTGTGAGTGGACATCGATTTATGTGAGCCTTAATCCCTGCCAGAACGGTGCCTCAGTGC	2033
QY	1621	TACAAACGTGCGAGTGAATTTCTGCAAGTGTGCCCCGAGGACTATGAGGGCAAGAACTGC	1680
DB	2034	TACAAACGTGCGAGTGAATTTCTGCAAGTGTGCCCCGAGGACTATGAGGGCAAGAACTGC	2093
QY	1681	TCACACCTGAAAGACCACTGCGCAGCACGCCCTGTGAAAGTGAATGACGTGCACAGTG	1740
DB	2094	TCACACCTGAAAGACCACTGCGCAGCACGCCCTGTGAAAGTGAATGACGTGCACAGTG	2153
QY	1741	GCCATGGCTTCCAAACGACACACCTGAAAGGGGTGCGGTATATTTCTCCAAAGTCTGTGGT	1800
DB	2154	GCCATGGCTTCCAAACGACACACCTGAAAGGGGTGCGGTATATTTCTCCAAAGTCTGTGGT	2213
QY	1801	CCTCACGGGAAGTCAAGAGTCACTGCGGGAGGGCAAAATTCACCTGTGACTGTFAACAAAGGC	1860
DB	2214	CCTCACGGGAAGTCAAGAGTCACTGCGGGAGGGCAAAATTCACCTGTGACTGTFAACAAAGGC	2273
QY	1861	TTTACGGGAACATCTGCGCAATTAATTAATGACTGTGAGAGCAACCCCTGTGAGAAC	1920
DB	2274	TTTACGGGAACATCTGCGCAATTAATTAATGACTGTGAGAGCAACCCCTGTGAGAAC	2333
QY	1921	GGTGGCACTTGCATCGATGGTCAACTCTCAAGTGCATCTGTAGTGAAGGCTGGGAG	1980
DB	2334	GGTGGCACTTGCATCGATGGTGTCAACTCTCAAGTGCATCTGTAGTGAAGGCTGGGAG	2393
QY	1981	GGGGCTTCTGTGAAACCAATATTAATGACTGTGAGGCAAGAACCCCTGCGCAAAATGGGGGC	2040
DB	2394	GGGGCTTCTGTGAAACCAATATTAATGACTGTGAGGCAAGAACCCCTGCGCAAAATGGGGGC	2453

QY	2041	ACGTGTGCGACCTGCTCAATGACTTCTACTGTGACTGTAAATAATGGGTGGAAAGAAAG	2100
DB	2454	ACGTGTGCGACCTGCTCAATGACTTCTACTGTGACTGTAAATAATGGGTGGAAAGAAAG	2513
QY	2101	ACCTGCCACTCAGCTGACAGTCACTGATGATGAGGCCACGTGCAACAAACGGTGGCACCTGC	2160
DB	2514	ACCTGCCACTCAGCTGACAGTCACTGATGATGAGGCCACGTGCAACAAACGGTGGCACCTGC	2573
QY	2161	TATGATGAGGGGAGTCTTTTAAAGTCACTGTCTCTGGCGGCTGGGAAAGAAACAACTGT	2220
DB	2574	TATGATGAGGGGAGTCTTTTAAAGTCACTGTCTCTGGCGGCTGGGAAAGAAACAACTGT	2633
QY	2221	AACATAGCCCGAAACAGTAGCTCCCTGCCAAACCCCTGCCATTAATGGGGGCAATGTGTG	2280
DB	2634	AACATAGCCCGAAACAGTAGCTCCCTGCCAAACCCCTGCCATTAATGGGGGCAATGTGTG	2693
QY	2281	GTCAACGGGAGTCTTTTACGTGCTGTGAGAGGAAGGCTGGGAGGGGCCCATCTGTGCT	2340
DB	2694	GTCAACGGGAGTCTTTTACGTGCTGTGAGAGGAAGGCTGGGAGGGGCCCATCTGTGCT	2753
QY	2341	CAGAATACCAATGACTGACGCCCTCATCCCTGTTTACAACAGCGGCACCTGTGTGGATGGA	2400
DB	2754	CAGAATACCAATGACTGACGCCCTCATCCCTGTTTACAACAGCGGCACCTGTGTGGATGGA	2813
QY	2401	GACAACTGGTACCGGTGCGAAATGTGCCCGGGTTTTGTGGGCCCGGACTGACAGAAATAAC	2460
DB	2814	GACAACTGGTACCGGTGCGAAATGTGCCCGGGTTTTGTGGGCCCGGACTGACAGAAATAAC	2873
QY	2461	ATCAATGAATGCGCAGTCTTCACTTGTGCTTTTGGAGCGACCTGTGTGGATGAGATCAAT	2520
DB	2874	ATCAATGAATGCGCAGTCTTCACTTGTGCTTTTGGAGCGACCTGTGTGGATGAGATCAAT	2933
QY	2521	GGCTACCGGTGTCTGCTCCAGGGCACAGTGTGCCAAGTGCAGGAAGTTTCAGGG	2580
DB	2934	GGCTACCGGTGTCTGCTCCAGGGCACAGTGTGCCAAGTGCAGGAAGTTTCAGGG	2993
QY	2581	AGAACCTTGATCAACATGGGGAGTGTATACCAAGATGGGGCCAAATGGGATGATGACTGT	2640
DB	2994	AGAACCTTGATCAACATGGGGAGTGTATACCAAGATGGGGCCAAATGGGATGATGACTGT	3053
QY	2641	AATACCTGCGAGTCTGTAATGACCGATGCTGCTCAAGGTCTGTGTGGTGGCCCTCGA	2700
DB	3054	AATACCTGCGAGTCTGTAATGACCGATGCTGCTCAAGGTCTGTGTGGTGGCCCTCGA	3113
QY	2701	CCTTGCTGTCTCCAAAGGGCACAGCGAGTGCCTCCAGCGGCGAGCTGCATCCCCATC	2760
DB	3114	CCTTGCTGTCTCCAAAGGGCACAGCGAGTGCCTCCAGCGGCGAGCTGCATCCCCATC	3173
QY	2761	CTGGACGACCGAGTGTCTGCTCCACCCCTGCACTGGTGTGGGGGAGTGTCCAGT	2820
DB	3174	CTGGACGACCGAGTGTCTGCTCCACCCCTGCACTGGTGTGGGGGAGTGTCCAGT	3233
QY	2821	CTCCAGCGGTGAAAGCAAAAGTGCACCTCTGACTCTTATTAACAGGATATCTGTGCGAAC	2880
DB	3234	CTCCAGCGGTGAAAGCAAAAGTGCACCTCTGACTCTTATTAACAGGATATCTGTGCGAAC	3293
QY	2881	ATCACATTTTACCTTACAGAGGAGATGATCTCACAGGTCTTACTAGCGAGCAATTTGC	2940
DB	3294	ATCACATTTTACCTTACAGAGGAGATGATCTCACAGGTCTTACTAGCGAGCAATTTGC	3353
QY	2941	AGTGAATTTGAGGAATTTGAATATTTTGAAGAATTTTCCGCTGAATATTCATCTACATC	3000
DB	3354	AGTGAATTTGAGGAATTTGAATATTTTGAAGAATTTTCCGCTGAATATTCATCTACATC	3413
QY	3001	GCTTGCAGCTTCCCTTTCAGCGAAACAATGAATAATCATGTGGCCATTTCTGTGGAAGAT	3060
DB	3414	GCTTGCAGCTTCCCTTTCAGCGAAACAATGAATAATCATGTGGCCATTTCTGTGGAAGAT	3473
QY	3061	ATACGGGATGATGGGAACCGATCAAGGAAATCACTGCAAAAAATAATTCGATCTTGTGTAGT	3120
DB	3474	ATACGGGATGATGGGAACCGATCAAGGAAATCACTGCAAAAAATAATTCGATCTTGTGTAGT	3533
QY	3121	AAACGTGATGGAAACAGCTCGCTGATTTGCTGCCGTTGCAAGAGTGAAGTTTCAGAGGCGG	3180

```
Db 3534 AAACGCTGATGAAACAGCTCGCTGATTCCTCCCTTGCGAAGTAAGAGTTTCAGAGCGG 3593
Qy CCTCTGAAGAACAGAGAGATTTCTCTTGTTCCTTGTCTGAGCTCTGTCTTAACTTGTGGCT 3240
Db CCTCTGAAGAACAGAGATTTCTCTTGTTCCTTGTCTGAGCTCTGTCTTAACTTGTGGCT 3653
Qy TGGATCTGTGCTTGGTGACGGCCCTTCTACTGTGTGCTCGGGAAGCGCGGAGCCGGGC 3300
Db TGGATCTGTGCTTGGTGACGGCCCTTCTACTGTGTGCTCGGGAAGCGCGGAGCCGGGC 3713
Qy AGCCACACACACTCAGCTCTGAGGACAAACACACCAACAGCTGCGGAGCAGCTGAAC 3360
Db AGCCACACACACTCAGCTCTGAGGACAAACACACCAACAGCTGCGGAGCAGCTGAAC 3773
Qy CAGATCAAAAACCCCATTTGAGAAACATGGGGCCAAACACGGTCCCATCAAGGATTACGAG 3420
Db CAGATCAAAAACCCCATTTGAGAAACATGGGGCCAAACACGGTCCCATCAAGGATTACGAG 3833
Qy AACAGAACTCCAAATGTCTAAATTAAGGACACACAAATTTGAAGTAGAAGAGGACGAC 3480
Db AACAGAACTCCAAATGTCTAAATTAAGGACACACAAATTTGAAGTAGAAGAGGACGAC 3893
Qy ATGACAAAACACCAAGAGAGCCCGGTTTGGCAAGCAGCGCGGTATACGCTGTAGAC 3540
Db ATGACAAAACACCAAGAGAGCCCGGTTTGGCAAGCAGCGCGGTATACGCTGTAGAC 3953
Qy AGAGAAGAGAGCCCGCCCAAGCGGACGCCGACCAAAACACCCAACTGGACAAACACAG 3600
Db AGAGAAGAGAGCCCGCCCAAGCGGACGCCGACCAAAACACCCAACTGGACAAACACAG 4013
Qy GACAAACAGAGCTTGGAAAGTCCAGAGCTTAAACCGAATGGAGTACATCGTATAG 3657
Db GACAAACAGAGCTTGGAAAGTCCAGAGCTTAAACCGAATGGAGTACATCGTATAG 4070

RESULT 7
LOCUS CQ894703 5896 bp DNA linear PAT 05-NOV-2004
DEFINITION Sequence 13 from Patent EP1471075.
ACCESSION CQ894703
VERSION CQ894703.1 GI:55467452
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Rosenthal,A.D., Pilarsky,C., Dahl,E., Specht,T., Bruemendorf,T.,
Lichtner,R., Staub,E., Roepcke,S. and Li,X.I.
TITLE Human nucleic acid sequences expressed in pancreatic carcinomas
JOURNAL Patent: Ep 1471075-A 13 27-OCT-2004;
Hinzmann, Bernd (DE); Rosenthal, Andre (DE); Pilarsky, Christian
(DE); Dahl, Edgar (DE); Specht, Thomas (DE); Lichtner, Rosemarie
(DE)

FEATURES
source Location/Qualifiers
1..5896
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.9%; Score 3654.2; DB 6; Length 5896;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3653; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGCGTTCCACGACGACRCGCGCGCTTCGGGCGCCCTTAAGCTTCCTGCTCGCCCTG 60
Db 414 ATGCGTTCCACGACGACRCGCGCGCTTCGGGCGCCCTTAAGCTTCCTGCTCGCCCTG 473
Qy 61 CTCTGTGCTTGGAGCAAGGTGTGTGGGGCTTCGGGTGAGTTGAGATCTG 120
```

```
Db 474 CTCTGTGCTTGGAGCAAGGTGTGTGGGGCTTCGGGTGAGTTGAGATCTG 533
Qy 121 TCCATGCAAGAACGTGAACCGGGAGCTGCAGAACGGGAACCTGTCGGCGCGCCGGAAC 180
Db 534 TCCATGCAAGAACGTGAACCGGGAGCTGCAGAACGGGAACCTGTCGGCGCGCCGGAAC 593
Qy 181 CCGGAGACCGCAAGTGCAACCGGAGCAGTGTGACACATCTTCAAAGTGTGCTCAAG 240
Db 594 CCGGAGACCGCAAGTGCAACCGGAGCAGTGTGACACATCTTCAAAGTGTGCTCAAG 653
Qy 241 GAGTATCAGTCCCGCGCTCAGCGCGGGGGCCCTGCGAGCTTCGGCTCAGGGTCCACGGCT 300
Db 654 GAGTATCAGTCCCGCGCTCAGCGCGGGGGCCCTGCGAGCTTCGGCTCAGGGTCCACGGCT 713
Qy 301 GTCAATCGGGGGCAACACCTTCAAGCCAGCCGCGGCAACGACCGCAACCGCATC 360
Db 714 GTCAATCGGGGGCAACACCTTCAAGCCAGCCGCGGCAACGACCGCAACCGCATC 773
Qy 361 GTGCTGCTTTTCAGTTTCGGCTGGCGAGGTCTTATACGTTGCTTGTGGAGGCGTGGAT 420
Db 774 GTGCTGCTTTTCAGTTTCGGCTGGCGAGGTCTTATACGTTGCTTGTGGAGGCGTGGAT 833
Qy 421 TCCAGTAAATGACACCGGTTCAACCTGACAGTATTATTGAAAAGGCTTCTCACTCGGGCATG 480
Db 834 TCCAGTAAATGACACCGGTTCAACCTGACAGTATTATTGAAAAGGCTTCTCACTCGGGCATG 893
Qy 481 ATCAACCCAGCGCGGAGTGGCAGAGCGCTGAAGCAGAACACGGCGTGTGCCACTTTGAG 540
Db 894 ATCAACCCAGCGCGGAGTGGCAGAGCGCTGAAGCAGAACACGGCGTGTGCCACTTTGAG 953
Qy 541 TATCAGATCCGCTGACCTGTGATGACTACTACTATGGCTTTGGCTGTGAATTAAGTTCTGC 600
Db 954 TATCAGATCCGCTGACCTGTGATGACTACTACTATGGCTTTGGCTGTGAATTAAGTTCTGC 1013
Qy 601 CGCCCCAGAGATGACTTCTTTTGACACTATGCTGTCAGCAGAAATGCAACAAACATTGC 660
Db 1014 CGCCCCAGAGATGACTTCTTTTGACACTATGCTGTCAGCAGAAATGCAACAAACATTGC 1073
Qy 661 ATGGAAGGCTGGATGGGCCCGCAATGTAAACAGAGCTATTTCGCCGACAAGGCTGCAGTCT 720
Db 1074 ATGGAAGGCTGGATGGGCCCGCAATGTAAACAGAGCTATTTCGCCGACAAGGCTGCAGTCT 1133
Qy 721 AAGCATGGGCTTTTGAAAATCTCCAGGTAAGTGCAGTGGCAGTAYGGCTGGCAAGGCTG 780
Db 1134 AAGCATGGGCTTTTGAAAATCTCCAGGTAAGTGCAGTGGCAGTAYGGCTGGCAAGGCTG 1193
Qy 781 TACTGTGATAAGTGATCCACACCCGGGATGCTCCACGCGCATCTCTAATGAGCCCTGG 840
Db 1194 TACTGTGATAAGTGATCCACACCCGGGATGCTCCACGCGCATCTCTAATGAGCCCTGG 1253
Qy 841 CAGTGGCTCTGTGAGACCAACTGGGGCGGCGAGCTCTGTGACAAAGATCTCAATTAAGTGT 900
Db 1254 CAGTGGCTCTGTGAGACCAACTGGGGCGGCGAGCTCTGTGACAAAGATCTCAATTAAGTGT 1313
Qy 901 GGAAGTCTATCAGCGGTGTCTCAACGGGGGAACTTTAGCAACACAGGCCCTGCAAAATAT 960
Db 1314 GGAAGTCTATCAGCGGTGTCTCAACGGGGGAACTTTAGCAACACAGGCCCTGCAAAATAT 1373
Qy 961 CAGTGTTCCTGCTGAGGGGTATTTCAGGACCCAACTGTGAAATTTGCTGAGACGCGCTGC 1020
Db 1374 CAGTGTTCCTGCTGAGGGGTATTTCAGGACCCAACTGTGAAATTTGCTGAGACGCGCTGC 1433
Qy 1021 CTCTCTGATCCCTGTCAACACAGAGGCGAGCTGTAAAGAGACCTCCCTCGGGCTTTGAGTGT 1080
Db 1434 CTCTCTGATCCCTGTCAACACAGAGGCGAGCTGTAAAGAGACCTCCCTCGGGCTTTGAGTGT 1493
Qy 1081 GAGTGTTCCTGAGTGGACCGGCCCAATGCTCTTACAAACATTTGATGATGCTGTTCTCCT 1140
Db 1494 GAGTGTTCCTGAGTGGACCGGCCCAATGCTCTTACAAACATTTGATGATGCTGTTCTCCT 1553
Qy 1141 AATAACTGTTCCACCGGGGCACTGCCAGGACCTGCTTAACGGATTTAAGTGTGCTGTC 1200
Db 1554 AATAACTGTTCCACCGGGGCACTGCCAGGACCTGCTTAACGGATTTAAGTGTGCTGTC 1613
```

Qy	1201	CCCCCAGTGGACTGGGAAAACGTGCGAGTTAGATGCAATGAATGTTGAGGCCAACAACCT	1260
Db	1614	CCCCCAGTGGACTGGGAAAACGTGCGAGTTAGATGCAATGAATGTTGAGGCCAACAACCT	1673
Qy	1261	TGTGTAAACGCCAAATCCTGTGAAGAACTCATTTAGAAATCTCATTTGCCAGCTACTTCC	1320
Db	1674	TGTGTAAACGCCAAATCCTGTGAAGAACTCATTTAGAAATCTCATTTGCCAGCTACTTCC	1733
Qy	1321	GGCTGGATGGGTGAGAAATGTGACATATAATTAATGACTGCTTGGCCAGCTGTGAGAAAT	1380
Db	1734	GGCTGGATGGGTGAGAAATGTGACATATAATTAATGACTGCTTGGCCAGCTGTGAGAAAT	1793
Qy	1381	GACGCTCTCTGTGGGATTTGGTTAAATGGTTATCGCTGTATCTGTCCACCTGGCTATGCA	1440
Db	1794	GACGCTCTCTGTGGGATTTGGTTAAATGGTTATCGCTGTATCTGTCCACCTGGCTATGCA	1853
Qy	1441	GGCGATCACTGTGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGGT	1500
Db	1854	GGCGATCACTGTGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGGT	1913
Qy	1501	CACCTGCAGAAATCAACAGATTCAGTGTCTGTGTGCCACTGTGTTCTCTGGAAC	1560
Db	1914	CACCTGCAGAAATCAACAGATTCAGTGTCTGTGTGCCACTGTGTTCTCTGGAAC	1973
Qy	1561	CTCTGTGAGTGGACATCGATTATTTGTGAGCCTAATCCCTGCCAGACGGTGGCCAGTGC	1620
Db	1974	CTCTGTGAGTGGACATCGATTATTTGTGAGCCTAATCCCTGCCAGACGGTGGCCAGTGC	2033
Qy	1621	TACAACCGTGCAGTACTATTCTTGCAGATGCGCCGAGGACTATGAGGGCAAGAACTGC	1680
Db	2034	TACAACCGTGCAGTACTATTCTGCAATGCGCCGAGGACTATGAGGGCAAGAACTGC	2093
Qy	1681	TCACACCTGAAAGACCACTGCGGCAACGCCCTGTGTGAAGTGAATGACGTGCACAGTG	1740
Db	2094	TCACACCTGAAAGACCACTGCGGCAACGCCCTGTGTGAAGTGAATGACGTGCACAGTG	2153
Qy	1741	GCCATGGCTTCCACGACACACCTGAAAGGGGTGGGTATATTTCCCTCAACGCTGTGGT	1800
Db	2154	GCCATGGCTTCCACGACACACCTGAAAGGGGTGGGTATATTTCCCTCAACGCTGTGGT	2213
Qy	1801	CCTCAGGGAAAGTCAAGAGTCAGTCGGGAGGCAAAATTCACCTGTGACTGTGAAACAAAGGC	1860
Db	2214	CCTCAGGGAAAGTCAAGAGTCAGTCGGGAGGCAAAATTCACCTGTGACTGTGAAACAAAGGC	2273
Qy	1861	TTCAACGGGAAACATCTGTCATGAAATATTAATGACTGTGAGAGCAACCTTGTAGAAAC	1920
Db	2274	TTCAACGGGAAACATCTGTCATGAAATATTAATGACTGTGAGAGCAACCTTGTAGAAAC	2333
Qy	1921	GGTGGCACTTCATCGATGGTGTCAATCCTCAAGTGCATCTGTAGTGGCTGGGAG	1980
Db	2334	GGTGGCACTTCATCGATGGTGTCAATCCTCAAGTGCATCTGTAGTGGCTGGGAG	2393
Qy	1981	GGGGCTACTGTGAAACCAATTAATGACTGCGAGGCAAGACCCCTGCCACATGGGGGC	2040
Db	2394	GGGGCTACTGTGAAACCAATTAATGACTGCGAGGCAAGACCCCTGCCACATGGGGGC	2453
Qy	2041	ACGTGTCCGACCTGGTCAATGACTTCTGACTGTGATGCTGTGATAAATGGGTGAAAGGAAAG	2100
Db	2454	ACGTGTCCGACCTGGTCAATGACTTCTGACTGTGATGCTGTGATAAATGGGTGAAAGGAAAG	2513
Qy	2101	ACCTGCCACTCACGTGACAGTCAGTGTGATGAGGCCACGTGCACACCGTGGCACCTGC	2160
Db	2514	ACCTGCCACTCACGTGACAGTCAGTGTGATGAGGCCACGTGCACACCGTGGCACCTGC	2573
Qy	2161	TATGATGAGGGGATGCTTTTAAAGTGCATGTGTCTCTGGCGCTGGGAGGAAACACCTGT	2220
Db	2574	TATGATGAGGGGATGCTTTTAAAGTGCATGTGTCTCTGGCGCTGGGAGGAAACACCTGT	2633
Qy	2221	AACATAGCCCGAAACAGTAGTGTGCTGCCAACCCCTGCGCAATATGGGGGACATGTGTG	2280
Db	2634	AACATAGCCCGAAACAGTAGTGTGCTGCCAACCCCTGCGCAATATATGGGGGACATGTGTG	2693

Qy	2281	GTCAACGGCGAGTCTTTTACGTGCGTCTGCAAGAAAGGTGGAGGGGCCCATCTGTGCT	2340
Db	2694	GTCAACGGCGAGTCTTTTACGTGCGTCTGCAAGAAAGGTGGAGGGGCCCATCTGTGCT	2753
Qy	2341	CAGAAATACCAATGACTGCGAGCCCTCATCTCTGTGTTACAAACGCGGACCTGTGTGGATGGA	2400
Db	2754	CAGAAATACCAATGACTGCGAGCCCTCATCTCTGTGTTACAAACGCGGACCTGTGTGGATGGA	2813
Qy	2401	GACAACTGTGTACCGGTGCGAAATGTGCCCCGGGTTTTGCTGGGCCCGACTCAGAAATAAAC	2460
Db	2814	GACAACTGTGTACCGGTGCGAAATGTGCCCCGGGTTTTGCTGGGCCCGACTCAGAAATAAAC	2873
Qy	2461	ATCAATGAATGCCAGTCTTTCACCTTGTGCTTTTGGAGCGACCTGTGTGGATGAGATCAAT	2520
Db	2874	ATCAATGAATGCCAGTCTTTCACCTTGTGCTTTTGGAGCGACCTGTGTGGATGAGATCAAT	2933
Qy	2521	GGCTACCGGTGTGTGCTTCCAGGCGACAGTGTGGTGCMAAGTGCMAAGAAAGTTTCAAGG	2580
Db	2934	GGCTACCGGTGTGTGCTTCCAGGCGACAGTGTGGTGCMAAGTGCMAAGAAAGTTTCAAGG	2993
Qy	2581	AGACCTTGCATCACATGGGAGTGTGATACAGATGGGGCCAAATGGGATGATGACTGT	2640
Db	2994	AGACCTTGCATCACATGGGAGTGTGATACAGATGGGGCCAAATGGGATGATGACTGT	3053
Qy	2641	AATACCTGCGCAGTGTGATGACAGATGCGCTGTCTCAAAGGTCTGTGTGGCCCTCGA	2700
Db	3054	AATACCTGCGCAGTGTGATGACAGATGCGCTGTCTCAAAGGTCTGTGTGGCCCTCGA	3113
Qy	2701	CCTTGCCTGTCTCAAAAGGGCAACAGGAGTGCCTGCGGGGAGAGCTGCATCCCATC	2760
Db	3114	CCTTGCCTGTCTCAAAAGGGCAACAGGAGTGCCTGCGGGGAGAGCTGCATCCCATC	3173
Qy	2761	CTGAGACAGAGTCTTCTGTCACCTGCACTGTGTGGGCGAGTGTCTTCCAGT	2820
Db	3174	CTGAGACAGAGTCTTCTGTCACCTGCACTGTGTGGGCGAGTGTCTTCCAGT	3233
Qy	2821	CTCAGCGGTGAAAGCAAAAGTGCACCTCTGACTCTCTATTACAGAGTAATCTGTGCGAAC	2880
Db	3234	CTCAGCGGTGAAAGCAAAAGTGCACCTCTGACTCTCTATTACAGAGTAATCTGTGCGAAC	3293
Qy	2881	ATCACATTTACCTTTTAAAGGAGATGATGTCACAGGTCTTACTACGGAGCAATTTGC	2940
Db	3294	ATCACATTTACCTTTTAAAGGAGATGATGTCACAGGTCTTACTACGGAGCAATTTGC	3353
Qy	2941	AGTGAATTTAGGAAATTTGAATATTTTGAAGAAATTTCCGCTGAATATCAATCTACATC	3000
Db	3354	AGTGAATTTAGGAAATTTGAATATTTTGAAGAAATTTCCGCTGAATATCAATCTACATC	3413
Qy	3001	GCTTGCAGGCTTCCCTTTCAGCGAAACAAATGAAATACATGTGGCCATTTCTGTGAAAGAT	3060
Db	3414	GCTTGCAGGCTTCCCTTTCAGCGAAACAAATGAAATACATGTGGCCATTTCTGTGAAAGAT	3473
Qy	3061	ATACGGGATGATGGGAACCCGATCAAGGAAATCACTGACAAAATAATATCGATCTTTGTAGT	3120
Db	3474	ATACGGGATGATGGGAACCCGATCAAGGAAATCACTGACAAAATAATATCGATCTTTGTAGT	3533
Qy	3121	AAAGGTGATGGAACACCTGCTGATTTGTCGGTTCAGAGTGAAGAGTTTCAAGGCGG	3180
Db	3534	AAAGGTGATGGAACACCTGCTGATTTGTCGGTTCAGAGTGAAGAGTTTCAAGGCGG	3593
Qy	3181	CCTCTGAAGAACGAAACAGATTTTCCCTTGTCTGCTGAGCTCTGTCTTAACTGTGGCT	3240
Db	3594	CCTCTGAAGAACGAAACAGATTTTCCCTTGTCTGCTGAGCTCTGTCTTAACTGTGGCT	3653
Qy	3241	TGGATCTGTGTGTGGTGACGGCCCTTCTACTGTGCTTGGGAGCGCGGAGCCCGGC	3300
Db	3654	TGGATCTGTGTGTGGTGACGGCCCTTCTACTGTGCTTGGGAGCGCGGAGCCCGGC	3713
Qy	3301	AGCCACACACCTCAGCCTCTGAGGACAAACACCAACAGCTGCGGGAGCAGCTGAAC	3360
Db	3714	AGCCACACACCTCAGCCTCTGAGGACAAACACCAACAGCTGCGGGAGCAGCTGAAC	3773
Qy	3361	CAGATCAAAAAACCCCAATTGAGAAAACATGGGGCCAAACGCGTCCCTCATCAAGGATTTACGAG	3420

Db 3774 CAGATCAAAACCCCAATGGAACAATGGGCGCAACACGGTCCCATCAAGGATTCAGAG 3833
Qy 3421 AACAAAGAACTCCAAAATGTTCTAAATTAAGGACACACAATTTCTGAAGTAGAAGAGGACGAC 3480
Db 3834 AACAAAGAACTCCAAAATGTTCTAAATTAAGGACACACAATTTCTGAAGTAGAAGAGGACGAC 3893
Qy 3481 ATGACAAACACACAGCAGAAAGCCGGTTTGGGAGACAGCCGGGCTATACGCTGGTAGAC 3540
Db 3894 ATGACAAACACACAGCAGAAAGCCGGTTTGGGAGACAGCCGGGCTATACGCTGGTAGAC 3953
Qy 3541 AGAGAAGAAAGCCCGCCCAAGCGGACGCGGACAAACACACCCAACTCGACAAACAAACAG 3600
Db 3954 AGAGAAGAAAGCCCGCCCAAGCGGACGCGGACAAACACACCCAACTCGACAAACAAACAG 4013
Qy 3601 GACAAACAGAGACTTGGAAAGTGGCCAGAGCTTAAACCGAATGGAGTAGACATCGTATAG 3657
Db 4014 GACAAACAGAGACTTGGAAAGTGGCCAGAGCTTAAACCGAATGGAGTAGACATCGTATAG 4070

RESULT 8
CQ981309 5896 bp DNA linear PAT 25-JAN-2005
LOCUS Sequence 164 from Patent EP1498424.
ACCESSION CQ981309
VERSION CQ981309.1 GI:58190568
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1
AUTHORS Rosenthal, A., Hermann, K., Heiden, E., Pilarsky, C., Bruemendorf, T.,
Staub, E., Roepcke, S., Mennerich, D., Kinnemann, H. and Li, X.
TITLE Human nucleic acid sequences from lung tumours
JOURNAL Patent: EP 1498424-A 164 19-JAN-2005;
Hinzmann, Bernd (DE); Hermann, Klaus (DE); Heiden, Esmeralda (DE);
Rosenthal, Andre (DE)
FEATURES
source 1. 5896
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 99.9%; Score 3654.2; DB 6; Length 5896;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3653; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGCTTTCCACACGACGCGCGCGCTCCGGCGCCCTTAAGCCCTCTGCTCGCCCTG 60
Db 414 ATGCTTTCCACACGACGCGCGCGCTCCGGCGCCCTTAAGCCCTCTGCTCGCCCTG 473
Qy 61 CTCTGCTCCCTGCGAGCAAGGTGTGTGGGCTCGGGTCAAGTTTCAAGTGTGATCTG 120
Db 474 CTCTGCTCCCTGCGAGCAAGGTGTGTGGGCTCGGGTCAAGTTTCAAGTGTGATCTG 533
Qy 121 TCCATGACAGACGTGAACGGGAGCTGCAGAACGGGAACCTGCTGGCGGCGCCGGAAC 180
Db 534 TCCATGACAGACGTGAACGGGAGCTGCAGAACGGGAACCTGCTGGCGGCGCCGGAAC 593
Qy 181 CCGGGAGACCGAGTGCACCCGCGACGAGTGTGACATATCTCAAAAGTGTGCTCAAG 240
Db 594 CCGGGAGACCGAGTGCACCCGCGACGAGTGTGACATATCTCAAAAGTGTGCTCAAG 653
Qy 241 GAGTATCAGTCCCGCTCAAGCGCGGGGGCCCTGACAGTTCGGCTCAGGGTCAAGCCT 300
Db 654 GAGTATCAGTCCCGCTCAAGCGCGGGGGCCCTGACAGTTCGGCTCAGGGTCAAGCCT 713
Qy* 301 GTCATCGGGGGACACCTTCAACCTCAAGCGGAGCGCGGACCGACCGACCGCATC 360
Db 714 GTCATCGGGGGACACCTTCAACCTCAAGCGGAGCGCGGACCGACCGCATC 773

Qy 361 GTGCTGCTTTTCAAGTTTGGCTTGGCCGAGGTCCTATACGTTGTGTGGAGGCGTGGAT 420
Db 774 GTGCTGCTTTTCAAGTTTGGCTTGGCCGAGGTCCTATACGTTGTGTGGAGGCGTGGAT 833
Qy 421 TCCAGTAATGACACCGTTCAACCTGACAGTATTTTGAAGAGGCTTCTCACTCGGGCATG 480
Db 834 TCCAGTAATGACACCGTTCAACCTGACAGTATTTTGAAGAGGCTTCTCACTCGGGCATG 893
Qy 481 ATCAACCCCAAGCGGCGAGTGGCAGACGCTGAAGCAGAACACGGGCTTGGCCACTTTGAG 540
Db 894 ATCAACCCCAAGCGGCGAGTGGCAGACGCTGAAGCAGAACACGGGCTTGGCCACTTTGAG 953
Qy 541 TATCAGATCCCGTGAACCTGTGATGACTACTATATGGCTTTGGCTGYATAATCTTGC 600
Db 954 TATCAGATCCCGTGAACCTGTGATGACTACTATATGGCTTTGGCTGYATAATCTTGC 1013
Qy 601 CGCCCCAGAGATGACTTCTTTTGGACACTATGCTGTGACAGATGCAACCAAACTTGC 660
Db 1014 CGCCCCAGAGATGACTTCTTTTGGACACTATGCTGTGACAGATGCAACCAAACTTGC 1073
Qy 661 ATGGAAGGCTGGATGGGCCCCCGAATGTAAACAGAGCTATTTGCCGCAAGGCTGCAGTCT 720
Db 1074 ATGGAAGGCTGGATGGGCCCCCGAATGTAAACAGAGCTATTTGCCGCAAGGCTGCAGTCT 1133
Qy 721 AAGCATGGGTCTTGCAAAATCCCAAGGTGACTGCAAGTGGCCAGTAYGCTGGCAAGGCTG 780
Db 1134 AAGCATGGGTCTTGCAAAATCCCAAGGTGACTGCAAGTGGCCAGTAYGCTGGCAAGGCTG 1193
Qy 781 TACTGTGATAAGTGCATCCCAACCCGGGATGCTGCACGGCATCTGTATAGCCCTG 840
Db 1194 TACTGTGATAAGTGCATCCCAACCCGGGATGCTGCACGGCATCTGTATAGCCCTG 1253
Qy 841 CAGTGCTCTGTGAGACCAACTGGGCGGCGAGCTCTGTGACAAAGATCTCAATTTACTGT 900
Db 1254 CAGTGCTCTGTGAGACCAACTGGGCGGCGAGCTCTGTGACAAAGATCTCAATTTACTGT 1313
Qy 901 GGGACTCATCAGCGGTGTCTCAACGGGGAACTTTGTAGCAACACAGGCCCTTGACAAATAT 960
Db 1314 GGGACTCATCAGCGGTGTCTCAACGGGGAACTTTGTAGCAACACAGGCCCTTGACAAATAT 1373
Qy 961 CAGTGCTCTGCTGAGGGGTATTTCAGGACCCCACTGTGAAATTTGTGAGCAGCCCTG 1020
Db 1374 CAGTGCTCTGCTGAGGGGTATTTCAGGACCCCACTGTGAAATTTGTGAGCAGCCCTG 1433
Qy 1021 CTCTGTATCCCTGTCAACACAGAGCAGCTGTAAAGGAGACCTCCCTGGGCTTTGAGTGT 1080
Db 1434 CTCTGTATCCCTGTCAACACAGAGCAGCTGTAAAGGAGACCTCCCTGGGCTTTGAGTGT 1493
Qy 1081 GAGTGTTCCTCCAGGCTGGACCGGCGCCCACTGTCTCAAAACATTTGATGCTGTCTCCT 1140
Db 1494 GAGTGTTCCTCCAGGCTGGACCGGCGCCCACTGTCTCAAAACATTTGATGCTGTCTCCT 1553
Qy 1141 AATAACTGTTCCACGGGGCACTTGCAGGACCTGGTTAAACGATTTAAAGTGTGTGTC 1200
Db 1554 AATAACTGTTCCACGGGGCACTTGCAGGACCTGGTTAAACGATTTAAAGTGTGTGTC 1613
Qy 1201 CCCCCACAGTGGACTGGGAAACGTGCAGTTAGATGCAATGAACTGTGAGGCCCAACCT 1260
Db 1614 CCCCCACAGTGGACTGGGAAACGTGCAGTTAGATGCAATGAACTGTGAGGCCCAACCT 1673
Qy 1261 TGTGTAAACGCCAAATCTCTGTAAGAATCTCATTTGCCAGCTACTATGCGACTGTCTTCCC 1320
Db 1674 TGTGTAAACGCCAAATCTCTGTAAGAATCTCATTTGCCAGCTACTATGCGACTGTCTTCCC 1733
Qy 1321 GGTGTGATGGGTGAGATTTGTGACATAAATATTAAATGATGTCCTTGGCCAGTGTGAGAT 1380
Db 1734 GGTGTGATGGGTGAGATTTGTGACATAAATATTAAATGATGTCCTTGGCCAGTGTGAGAT 1793
Qy 1381 GAGCGCTCTGTGCGGATTTGGTTAATGGTTATCGCTGTATCTGTCACCTGGCTATGCA 1440
Db 1794 GAGCGCTCTGTGCGGATTTGGTTAATGGTTATCGCTGTATCTGTCACCTGGCTATGCA 1853

QY 1441 GGCATCACTGTGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGT 1500
Db |||||
QY 1854 GGCATCACTGTGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGT 1913
Db |||||
QY 1501 CACTGTGAGAAATCAACAGATCCAGAGTCTGTGTGCCACTGTGTTCTCTGGAAAC 1560
Db |||||
QY 1914 CACTGTGAGAAATCAACAGATCCAGAGTCTGTGTGCCACTGTGTTCTCTGGAAAC 1973
QY |||||
QY 1561 CTCTGTGAGTGAACATCGATTTATGTAGCCTTAATCCCTGCCAGAACCGTGCAGTGC 1620
Db |||||
QY 1974 CTCTGTGAGTGAACATCGATTTATGTAGCCTTAATCCCTGCCAGAACCGTGCAGTGC 2033
QY |||||
QY 1621 TACAACCGTCCAGTGAATTTCTGCAAGTGCCTCCAGAGCTATGAGGCAAGAACTGC 1680
Db |||||
QY 2034 TACAACCGTCCAGTGAATTTCTGCAAGTGCCTCCAGAGCTATGAGGCAAGAACTGC 2093
QY |||||
QY 1681 TCACACCTGAAAGACCACTGCGGCAAGACCCCTGTGAAAGTGAATGACGTGCACAGTG 1740
Db |||||
QY 2094 TCACACCTGAAAGACCACTGCGGCAAGACCCCTGTGAAAGTGAATGACGTGCACAGTG 2153
QY |||||
QY 1741 GCCATGGTTCCAAAGACACACCTGAAAGGGGTGCGGTATATTTCTCCAAAGTCTGTGGT 1800
Db |||||
QY 2154 GCCATGGTTCCAAAGACACACCTGAAAGGGGTGCGGTATATTTCTCCAAAGTCTGTGGT 2213
QY |||||
QY 1801 CCTCACGGGAGTCAAGAGTCACTGCGGAGGCAAAATTCACCTGTGTGACTGTAAACAAAGGC 1860
Db |||||
QY 2214 CCTCACGGGAGTCAAGAGTCACTGCGGAGGCAAAATTCACCTGTGTGACTGTAAACAAAGGC 2273
QY |||||
QY 1861 TTCAACGGGAAACATCTGCCATGAAATAATTAATGACTGTGAGAGCAACCCCTGTGAGAAAC 1920
Db |||||
QY 2274 TTCAACGGGAAACATCTGCCATGAAATAATTAATGACTGTGAGAGCAACCCCTGTGAGAAAC 2333
QY |||||
QY 1921 GGTGGCACTTCGATCGATGTTGTCATCTCTCAAGTGCATCTGTAGTGAACGCTGGAG 1980
Db |||||
QY 2334 GGTGGCACTTCGATCGATGTTGTCATCTCTCAAGTGCATCTGTAGTGAACGCTGGAG 2393
QY |||||
QY 1981 GGGCCCTACTGTGAAACCAATTAATGACTGTGAGGCAAGCCCTGCCAATGGGGGC 2040
Db |||||
QY 2394 GGGCCCTACTGTGAAACCAATTAATGACTGTGAGGCAAGCCCTGCCAATGGGGGC 2453
QY |||||
QY 2041 ACCTGTGCGGACCTGGTCAATGACTTCTACTGTGACTGTAAATAATGGGTGAAAGGAAAG 2100
Db |||||
QY 2454 ACCTGTGCGGACCTGGTCAATGACTTCTACTGTGACTGTAAATAATGGGTGAAAGGAAAG 2513
QY |||||
QY 2101 ACCTGCCACTCAAGTGAACGTCAGTGTGAATGAGGCAAGTGCACAAACGCTGGACCTGC 2160
Db |||||
QY 2514 ACCTGCCACTCAAGTGAACGTCAGTGTGAATGAGGCAAGTGCACAAACGCTGGACCTGC 2573
QY |||||
QY 2161 TATGATGAGGGGGATGCTTTTAAAGTGCATGTCTCTGCGGCTGGGAAGGAACAACTGT 2220
Db |||||
QY 2574 TATGATGAGGGGGATGCTTTTAAAGTGCATGTCTCTGCGGCTGGGAAGGAACAACTGT 2633
QY |||||
QY 2221 AACATAGCCGAAACAGTAGTGCCTGCCAACCCCTGCCAATAATGGGGGCACATGTGTG 2280
Db |||||
QY 2634 AACATAGCCGAAACAGTAGTGCCTGCCAACCCCTGCCAATAATGGGGGCACATGTGTG 2693
QY |||||
QY 2281 GTCAACGCGAGTCTTTACGTGGGTCTGCAAGAGAGCTGGGAGGGGCCCATCTGTGCT 2340
Db |||||
QY 2694 GTCAACGCGAGTCTTTACGTGGGTCTGCAAGAGAGCTGGGAGGGGCCCATCTGTGCT 2753
QY |||||
QY 2341 CAGAAATCAATGACTGAGCGCTCATCTCTGTTTACAAACAGCGGCACCTGTGTGGATGGA 2400
Db |||||
QY 2754 CAGAAATCAATGACTGAGCGCTCATCTCTGTTTACAAACAGCGGCACCTGTGTGGATGGA 2813
QY |||||
QY 2401 GACAACTGGTACCGGTGGAATGTGCCCGGGTTTTGCTGGGGCCGCACTGCAGAAATAAC 2460
Db |||||
QY 2814 GACAACTGGTACCGGTGGAATGTGCCCGGGTTTTGCTGGGGCCGCACTGCAGAAATAAC 2873
QY |||||
QY 2461 ATCAATGAATGCCAGTCTTCACTTGTGCTTTGGAGGCACTGTGTGGATGAGATCAAT 2520
Db |||||
QY 2874 ATCAATGAATGCCAGTCTTCACTTGTGCTTTGGAGGCACTGTGTGGATGAGATCAAT 2933
QY |||||
QY 2521 GGCTACCGGTGTGTCTGCGCTCCAGGGCACAGTGTGTGCGCAAGTGCAGGAAGTTTCAGGG 2580

Db 2934 GGCTACCGGTGTGTCTGCGCTCCAGGGCACAGTGTGTGCCAAAGTGCAGGAAGTTTCAGGG 2993
QY |||||
QY 2581 AGACCTTGATCAACATGCGGAGTGTGATACAGATGGGGCCAAATGGGATGATGACTGT 2640
Db |||||
QY 2994 AGACCTTGATCAACATGCGGAGTGTGATACAGATGGGGCCAAATGGGATGATGACTGT 3053
QY |||||
QY 2641 AATACCTGCCAGTCCCTGAATGGAGCGATCGCTGCTCAAAAGTCTGGTGTGGCCCTCGA 2700
Db |||||
QY 3054 AATACCTGCCAGTCCCTGAATGGAGCGATCGCTGCTCAAAAGTCTGGTGTGGCCCTCGA 3113
QY |||||
QY 2701 CCTTGCTGTCTCCACAAAGGGCACAGCGAGTGCCTCCAGCGGGCAGAGCTGCATCCCATC 2760
Db |||||
QY 3114 CCTTGCTGTCTCCACAAAGGGCACAGCGAGTGCCTCCAGCGGGCAGAGCTGCATCCCATC 3173
QY |||||
QY 2761 CTGGACGACCGAGTCTTCTGTCACCCCTGCACTGTGTGTGGGCGAGTGTGGGTCTTCAGT 2820
Db |||||
QY 3174 CTGGACGACCGAGTCTTCTGTCACCCCTGCACTGTGTGTGGGCGAGTGTGGGTCTTCAGT 3233
QY |||||
QY 2821 CTCACGCGGTGAAGCAAAAGTGCACCTCTGACTCTCTATTACAGGATTAATCTGCGAAC 2880
Db |||||
QY 3234 CTCACGCGGTGAAGCAAAAGTGCACCTCTGACTCTCTATTACAGGATTAATCTGCGAAC 3293
QY |||||
QY 2881 ATCACAATTTACCTTTTAAACAGGAGATGATGTCAACAGGTCTTACTACGGAGCAATTTGC 2940
Db |||||
QY 3294 ATCACAATTTACCTTTTAAACAGGAGATGATGTCAACAGGTCTTACTACGGAGCAATTTGC 3353
QY |||||
QY 2941 AGTGAATTTAGGAATTTGAATATTTTGAAGATGTTTTCGCTGAATATTTCAATCTACATC 3000
Db |||||
QY 3354 AGTGAATTTAGGAATTTGAATATTTTGAAGATGTTTTCGCTGAATATTTCAATCTACATC 3413
QY |||||
QY 3001 GCTTGGCGGCTTCCCTTCAGCGAACAAATGAAATACATGTGCCCATTTCTGCTGAAAGAT 3060
Db |||||
QY 3414 GCTTGGCGGCTTCCCTTCAGCGAACAAATGAAATACATGTGCCCATTTCTGCTGAAAGAT 3473
QY |||||
QY 3061 ATACGGGATGATGGGAACCCGATCAAGGAAATCACTGACAAAAATAATTCGATCTTGTAGT 3120
Db |||||
QY 3474 ATACGGGATGATGGGAACCCGATCAAGGAAATCACTGACAAAAATAATTCGATCTTGTAGT 3533
QY |||||
QY 3121 AAACGTGATGGAACAGCTCGCTGATGTCGCGGTCGAGAGTAAGAGTTCAGAGCGG 3180
Db |||||
QY 3534 AAACGTGATGGAACAGCTCGCTGATGTCGCGGTCGAGAGTAAGAGTTCAGAGCGG 3593
QY |||||
QY 3181 CCTCTGAAGAACAGAACAGATTTCTTGTTCCTCTGCTGAGCTCTGTCTTAACTGTGGCT 3240
Db |||||
QY 3594 CCTCTGAAGAACAGAACAGATTTCTTGTTCCTCTGCTGAGCTCTGTCTTAACTGTGGCT 3653
QY |||||
QY 3241 TGGATCTGTTGCTTGGTGAACGCGCTTCTACTGCTGCTGCGGAAGCGGCGGAACCGGCG 3300
Db |||||
QY 3654 TGGATCTGTTGCTTGGTGAACGCGCTTCTACTGCTGCTGCGGAAGCGGCGGAACCGGCG 3713
QY |||||
QY 3301 AGCCACACACTCTGAGGACAAACACCAACCAACAGCTGCGGGAGCAGCTGAC 3360
Db |||||
QY 3714 AGCCACACACTCTGAGGACAAACACCAACCAACAGCTGCGGGAGCAGCTGAC 3773
QY |||||
QY 3361 CAGATCAAAACCCCTTGAAGAACATGGGCGCAACACGCTGCCCATCAAGGATTAACAG 3420
Db |||||
QY 3774 CAGATCAAAACCCCTTGAAGAACATGGGCGCAACACGCTGCCCATCAAGGATTAACAG 3833
QY |||||
QY 3421 AACAAAGAACTCCAAATGTCTTAAATAAGGACACAAATTTCTGAAGTAGAAGAGGACGAC 3480
Db |||||
QY 3834 AACAAAGAACTCCAAATGTCTTAAATAAGGACACAAATTTCTGAAGTAGAAGAGGACGAC 3893
QY |||||
QY 3481 ATGGAACAAACACCAAGGAAAGCCCGTTTGGCAGCAGCGGGGTATACGCTGGTAGAC 3540
Db |||||
QY 3894 ATGGAACAAACACCAAGGAAAGCCCGTTTGGCAGCAGCGGGGTATACGCTGGTAGAC 3953
QY |||||
QY 3541 AGAAGAGAGAGCCCGCCCAACCGCAGCAACCAACCAACTGCAACAAACAG 3600
Db |||||
QY 3954 AGAAGAGAGAGCCCGCCCAACCGCAGCAACCAACCAACTGCAACAAACAG 4013
QY |||||
QY 3601 GACAAACAGAGACTTGGAAAGTGCAGAGCTTAAACCGAAATGGAGTATCATCGTATAG 3657

Db	4014	GACACAGAGACTTGGAAAGTCCCGACAGCTTAAACCGAATGGAGTACATCGTATAG	4070
Db	1014	CGCCCCAGAGATGACTTCTTTGGACACTATGCTGTGACCAAGATGGCAACAAACTTGC	1073
Qy	661	ATGGAAGGCTGGATGGGCCCCGAAATGTAACAGAGCTATTTTGCCGACAAGGCTGCAGTCC	720
Db	1074	ATGGAAGGCTGGATGGGCCCCGAAATGTAACAGAGCTATTTTGCCGACAAGGCTGCAGTCC	1133
Qy	721	AAGATGGGTCTTGCAAACTCCACAGGTGATCGAGGTGCCAGTAYGGCTGGCAAGGCTG	780
Db	1134	AAGATGGGTCTTGCAAACTCCACAGGTGATCGAGGTGCCAGTAYGGCTGGCAAGGCTG	1193
Qy	781	TACTGTGATAAGTGCATCCACACCGGGATGCGTCCACGGCATCTGTAAATGAGCCCTGG	840
Db	1194	TACTGTGATAAGTGCATCCACACCGGGATGCGTCCACGGCATCTGTAAATGAGCCCTGG	1253
Qy	841	CAGTGCTCTGTGAGAACCAACTGGGGCGGCCAGCTCTGTGACAAAGATCTCAATTTACTGT	900
Db	1254	CAGTGCTCTGTGAGAACCAACTGGGGCGGCCAGCTCTGTGACAAAGATCTCAATTTACTGT	1313
Qy	901	GGGACTCATCAGCCGTGTCTCAACGGGGGAACTTTGAGCAACACAGAGCCCTGACAAATAT	960
Db	1314	GGGACTCATCAGCCGTGTCTCAACGGGGGAACTTTGAGCAACACAGAGCCCTGACAAATAT	1373
Qy	961	CAGTGTTCTGCTGCGCTGAGGGGTATTTACGACCCCACTGTGAAATTTGCTGACACGCTGC	1020
Db	1374	CAGTGTTCTGCTGCGCTGAGGGGTATTTACGACCCCACTGTGAAATTTGCTGACACGCTGC	1433
Qy	1021	CTCTCTGATCCCTGTCTCAACAGAGGCGAGCTGTAAAGGAGACTCTCTCGGCTTTGAGTGT	1080
Db	1434	CTCTCTGATCCCTGTCTCAACAGAGGCGAGCTGTAAAGGAGACTCTCTCGGCTTTGAGTGT	1493
Qy	1081	GAGTGTTCCCAAGGCTGGACCGGCCCCACATGCTCTCAAAAATTTGATGACTGTCTCTCT	1140
Db	1494	GAGTGTTCCCAAGGCTGGACCGGCCCCACATGCTCTCAAAAATTTGATGACTGTCTCTCT	1553
Qy	1141	AATACTGTTCCACGGGGGACCTGCCAGGACCTGGTTAAAGGATTTAAGTGTGTGTGC	1200
Db	1554	AATACTGTTCCACGGGGGACCTGCCAGGACCTGGTTAAAGGATTTAAGTGTGTGTGC	1613
Qy	1201	CCCCACAGTGACCTGGGAAAAAGCTGCCAGTTAGATGCAATGTAATGTGAGGCCAAACCT	1260
Db	1614	CCCCACAGTGACCTGGGAAAAAGCTGCCAGTTAGATGCAATGTAATGTGAGGCCAAACCT	1673
Qy	1261	TGTGTAAACCGCAAACTCTGTAAAGATCTCAATGCGAGCTACTACTGCGACTGTCTTCCC	1320
Db	1674	TGTGTAAACCGCAAACTCTGTAAAGATCTCAATGCGAGCTACTACTGCGACTGTCTTCCC	1733
Qy	1321	GGCTGATGGGTGAGATTTGACATAAATATTAATGATGCTGCTTGGCAGTGTGAGAT	1380
Db	1734	GGCTGATGGGTGAGATTTGACATAAATATTAATGATGCTGCTTGGCAGTGTGAGAT	1793
Qy	1381	GACGCTCTGTCGGGATTTGGTTAATGTTATCGTCTATCTGTCCACCTGGCTATGCA	1440
Db	1794	GACGCTCTGTCGGGATTTGGTTAATGTTATCGTCTATCTGTCCACCTGGCTATGCA	1853
Qy	1441	GGCGATCACTGTGAGAGAGACATCGATGAATGTGACGAAACCCCTGTTGAATGGGGGT	1500
Db	1854	GGCGATCACTGTGAGAGAGACATCGATGAATGTGACGAAACCCCTGTTGAATGGGGGT	1913
Qy	1501	CACTGTGAGAAATCAACAGATTTCCAGTGTCTGTGTCCCATCTGGTTTCTCTGGAAAC	1560
Db	1914	CACTGTGAGAAATCAACAGATTTCCAGTGTCTGTGTCCCATCTGGTTTCTCTGGAAAC	1973
Qy	1561	CTCTGTGAGTGCAGTATTTGTCAGCTTAATCCCTGCCAGAACGCTGCCAGTGC	1620
Db	1974	CTCTGTGAGTGCAGTATTTGTCAGCTTAATCCCTGCCAGAACGCTGCCAGTGC	2033
Qy	1621	TACAAACCTGTCAGTACTATTTCTGCAAGTGCCTGCCAGGACTATGAGGGCAAGAACTGC	1680
Db	2034	TACAAACCTGTCAGTACTATTTCTGCAAGTGCCTGCCAGGACTATGAGGGCAAGAACTGC	2093
Qy	1681	TCAACCTGAAAGACCACTGCGCAGACGACCCCTCTGTGAAGTGAATGACGTGACAGTG	1740

Db	2094	TCACACCTGAAAGACCACCTGCGGCACGACCCCTGTGAAAGTGAATTGACAGCTGCACAGTG	2153
Qy	1741	GCATGCGCTTCCACGACACACCTGAAGGGTGGGTATATTTCCTCCACAGCTGTGGT	1800
Db	2154	GCCATGCGCTTCCACGACACACCTGAAGGGGTGGGTATATTTCCTCCACAGCTGTGGT	2213
Qy	1801	CCTCACGGGAAGTGCAGAGTCACTCGGGAGGCAAAATTCACCTGTGACTGTGAATAAAGGC	1860
Db	2214	CCTCACGGGAAGTGCAGAGTCACTCGGGAGGCAAAATTCACCTGTGACTGTGAATAAAGGC	2273
Qy	1861	TTACCGGGAACATCTGCCATGAATAATTAATGACTGTGAGAGCAACCCCTGTGAGAAC	1920
Db	2274	TTACCGGGAACATCTGCCATGAATAATTAATGACTGTGAGAGCAACCCCTGTGAGAAC	2333
Qy	1921	GGTGGCACTTGCACTGATGGTGTCAACTCTACAGTGCATCTGTAGTGAAGGCTGGAG	1980
Db	2334	GGTGGCACTTGCACTGATGGTGTCAACTCTACAGTGCATCTGTAGTGAAGGCTGGAG	2393
Qy	1981	GGGGCCTACTGTGAACCAATATTAATGACTGACGCCAGAACCCCTGCCACAATGGGGGC	2040
Db	2394	GGGGCCTACTGTGAACCAATATTAATGACTGACGCCAGAACCCCTGCCACAATGGGGGC	2453
Qy	2041	ACGTGTCGCGACCTGGTCAATGACTTCTACTGTGACTGTAAAAATGGGTGAAAGGAAAG	2100
Db	2454	ACGTGTCGCGACCTGGTCAATGACTTCTACTGTGACTGTAAAAATGGGTGAAAGGAAAG	2513
Qy	2101	ACCTGCCACTCACGTGACAGTCACTGTGAGCGCAGTGCACAACAGGTGGCACCTGC	2160
Db	2514	ACCTGCCACTCACGTGACAGTCACTGTGAGCGCAGTGCACAACAGGTGGCACCTGC	2573
Qy	2161	TATCATGAGGGGATGCTTTTAAGTGCATGTGTCTGCGCGCTGGGAAGGAACAACCTGT	2220
Db	2574	TATCATGAGGGGATGCTTTTAAGTGCATGTGTCTGCGCGCTGGGAAGGAACAACCTGT	2633
Qy	2221	AACATAGCCGAAACAGTAGCTGCTGCCAACCCCTGCCAATATGGGGGCACATGTGTG	2280
Db	2634	AACATAGCCGAAACAGTAGCTGCTGCCAACCCCTGCCAATATGGGGGCACATGTGTG	2693
Qy	2281	GTCAACGCGAGTCTTTACGTGCGTCTGCAAGGAAGCTGGGAGGGGCCCATCTGTGCT	2340
Db	2694	GTCAACGCGAGTCTTTACGTGCGTCTGCAAGGAAGCTGGGAGGGGCCCATCTGTGCT	2753
Qy	2341	CAGAATACCAATGACTGCAGCGCTCATCCCTGTTTACAACAGCGGCACCTGTGTGGATGGA	2400
Db	2754	CAGAATACCAATGACTGCAGCGCTCATCCCTGTTTACAACAGCGGCACCTGTGTGGATGGA	2813
Qy	2401	GACAACTGGTACCGGTGCGAATGTGCCCGGGTTTTGCTGGGCCCGACTGCAGAAATAAC	2460
Db	2814	GACAACTGGTACCGGTGCGAATGTGCCCGGGTTTTGCTGGGCCCGACTGCAGAAATAAC	2873
Qy	2461	ATCAATGAATGCCAGTCTTTCACCTTGTGCTTTGGAGGGACCTGTGTGGATGAGATCAAT	2520
Db	2874	ATCAATGAATGCCAGTCTTTCACCTTGTGCTTTGGAGGGACCTGTGTGGATGAGATCAAT	2933
Qy	2521	GGCTACCGGTGTGTGCTTCCCTCCAGGCGACAGTGTGCAAGTGCAGGAAGTTTCAGGG	2580
Db	2934	GGCTACCGGTGTGTGCTTCCCTCCAGGCGACAGTGTGCAAGTGCAGGAAGTTTCAGGG	2993
Qy	2581	AGACCTTGCAATCACCATGGGAGTGTGTATACAGATGGGGCCAAATGGGATGATGACTGT	2640
Db	2994	AGACCTTGCAATCACCATGGGAGTGTGTATACAGATGGGGCCAAATGGGATGATGACTGT	3053
Qy	2641	AATACCTGCCAGTGTGTAATGGAGGATCGCTGCTCAAAAGGTGTGTTGGGCCCTCGA	2700
Db	3054	AATACCTGCCAGTGTGTAATGGAGGATCGCTGCTCAAAAGGTGTGTTGGGCCCTCGA	3113
Qy	2701	CCTTGCTGCTCCACAAAGGGCACAGCGAGTGGCCAGCGGCGAGCTGCATCCCCATC	2760
Db	3114	CCTTGCTGCTCCACAAAGGGCACAGCGAGTGGCCAGCGGCGAGCTGCATCCCCATC	3173
Qy	2761	CTGGACGACCAAGTGTTCCTGCACCCCTGCACTGGTGTGGCGAGTGTGCGTCTTCCAGT	2820
Db	3174	CTGGACGACCAAGTGTTCCTGCACCCCTGCACTGGTGTGGCGAGTGTGCGTCTTCCAGT	3233

Qy	2821	CTCCAGCCGTGAAGACAAAGTGCACCTCTGACTCCTCTATTACCAGGATAACTGTGGAAC	2880
Db	3234	CTCCAGCCGTGAAGACAAAGTGCACCTCTGACTCCTATTACCAGGATAACTGTGGAAC	3293
Qy	2881	ATCACATTTACCTTTAAACAGGAGATGATGTCAACAGGTCTTACTACGGAGCAATTTGC	2940
Db	3294	ATCACATTTACCTTTAAACAGGAGATGATGTCAACAGGTCTTACTACGGAGCAATTTGC	3353
Qy	2941	AGTGAATTCAGGAATTTGAATATTTTGAAGAATGTTTCCGCTGAATATTCAATCTACATC	3000
Db	3354	AGTGAATTCAGGAATTTGAATATTTTGAAGAATGTTTCCGCTGAATATTCAATCTACATC	3413
Qy	3001	GCTTTCGAGCCTTCCCTTCAGCGAAACAATGAATAATACATGTGGCCATTTCTGCTGAAGAT	3060
Db	3414	GCTTTCGAGCCTTCCCTTCAGCGAAACAATGAATAATACATGTGGCCATTTCTGCTGAAGAT	3473
Qy	3061	ATACGGGATGATGGGAACCCGATCAAGGAAATCACTGAACAAATAATCGATCTTGTGTAGT	3120
Db	3474	ATACGGGATGATGGGAACCCGATCAAGGAAATCACTGAACAAATAATCGATCTTGTGTAGT	3533
Qy	3121	AAACGTGATGGAACACAGCTCGCTGATTTGTCGCGTTCGAGAACTAAGAGTTCAGAGGCGG	3180
Db	3534	AAACGTGATGGAACACAGCTCGCTGATTTGTCGCGTTCGAGAACTAAGAGTTCAGAGGCGG	3593
Qy	3181	CCTCTGAAGAAACAGAAACAGATTTCTTGTTCCTTGTCTGAGCTCTGTCTTAACTGTGGCT	3240
Db	3594	CCTCTGAAGAAACAGAAACAGATTTCTTGTTCCTTGTCTGAGCTCTGTCTTAACTGTGGCT	3653
Qy	3241	TGGATCTGTTGTTGTTGAGCGGCTTCTACTGTGTGCTGCGGAAGCGGCGGAGCCGGGC	3300
Db	3654	TGGATCTGTTGTTGTTGAGCGGCTTCTACTGTGTGCTGCGGAAGCGGCGGAGCCGGGC	3713
Qy	3301	AGCCACACACACTCAGGCTCTGAGGACACACACCAACAACTGCGGGAGCAGCTGAAC	3360
Db	3714	AGCCACACACACTCAGGCTCTGAGGACACACACCAACAACTGCGGGAGCAGCTGAAC	3773
Qy	3361	CAGATCAAAAAACCCCAATTGAGAAACATGGGCGCAACACCGTCCCATCAAGGATTAAGAG	3420
Db	3774	CAGATCAAAAAACCCCAATTGAGAAACATGGGCGCAACACCGTCCCATCAAGGATTAAGAG	3833
Qy	3421	AACAAGAACTCCAAAAATGTCTAAAAATAAGACACACAATTTCTGAAGTAGAAGAGACAC	3480
Db	3834	AACAAGAACTCCAAAAATGTCTAAAAATAAGACACACAATTTCTGAAGTAGAAGAGACAC	3893
Qy	3481	ATGGACAAACACACGACGAGAAAGCCCGTTTGGCAGCAGCGCGGTATACGCTGGTAGAC	3540
Db	3894	ATGGACAAACACACGACGAGAAAGCCCGTTTGGCAGCAGCGCGGTATACGCTGGTAGAC	3953
Qy	3541	AGAGACAGAGAGCCGCCCAACCGCACCGCCGACAAAAACACCAAACTGGACAAAAACAAACAG	3600
Db	3954	AGAGACAGAGAGCCGCCCAACCGCACCGCCGACAAAAACACCAAACTGGACAAAAACAAACAG	4013
Qy	3601	GACACACAGAGACTTGGAAAGTCCCGAGAGCTTAAACCGAATGGAGTACATCGTATAG	3657
Db	4014	GACACACAGAGACTTGGAAAGTCCCGAGAGCTTAAACCGAATGGAGTACATCGTATAG	4070

RESULT 10
AX578038
LOCUS
DEFINITION Sequence 160 from Patent WO02081745.
ACCESSION AX578038
VERSION AX578038.1 GI:27647246
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Garcia,T., roman Roman,S., Baron,R., Call,K., Theillhaber,J., Connolly,T., Jackson,A., Bushnell,S.E. and Rawadi,G.

TITLE Genes involved in osteogenesis, and methods of use
JOURNAL Patent: WO 02081745-A 160 17-OCT-2002;
Aventis Pharma S.A. (FR)

FEATURES
source

Location/Qualifiers
1. 5896
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Homo sapiens jagged 1 (Alagille syndrome) (JAG1)"

ORIGIN

Query Match 99.9%; Score 3654.2; DB 6; Length 5896;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3653; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGCGTTCCCCACGACRCCGGCGGCTCCGGCGCCCTTAAGCCCTCTGCTCGCCCTG 60
Db 414 ATGCGTTCCCCACGACGCGCGCGGCTCCGGCGCCCTTAAGCCCTCTGCTCGCCCTG 473
Qy 61 CTCTGTGCTTCGAGCAAGGTGTGTGGGCTCTGGGTCAAGTTGAGATCTG 120
Db 474 CTCTGTGCTTCGAGCAAGGTGTGTGGGCTCTGGGTCAAGTTGAGATCTG 533
Qy 121 TCCATGACAGACGTGAACGGGAGCTGCAGAACGGGAACCTGCTCGCGCGGCCCGGAAC 180
Db 534 TCCATGACAGACGTGAACGGGAGCTGCAGAACGGGAACCTGCTCGCGCGGCCCGGAAC 593
Qy 181 CCGGGAGACCGCAAGTGCACCCGCGAGAGTGTGACACATCTTCAAAGTGTGCTCAAG 240
Db 594 CCGGGAGACCGCAAGTGCACCCGCGAGAGTGTGACACATCTTCAAAGTGTGCTCAAG 653
Qy 241 GAGTATCAGTCCCGCTCAAGCGCGGGGCTCTGAGCTTCGGCTCAGGGTCCAGCCT 300
Db 654 GAGTATCAGTCCCGCTCAGCGCGGGGCTCTGAGCTTCGGCTCAGGGTCCAGCCT 713
Qy 301 GTCATCGGGGCAACACTTCAACTCAAGCCAGCGCGGCAACGACCGCAACCGCATC 360
Db 714 GTCATCGGGGCAACACTTCAACTCAAGCCAGCGCGGCAACGACCGCAACCGCATC 773
Qy 361 GTGCTGCTTTTCAAGTTTGGCTGCGCGAGTCTTATACGTTGCTGTGGAGGGGTGGAT 420
Db 774 GTGCTGCTTTTCAAGTTTGGCTGCGCGAGTCTTATACGTTGCTGTGGAGGGGTGGAT 833
Qy 421 TCCAGTAATGACACCGTTCAACCTGACAGTATTAATGAAAGGCTTCTCACTCGGGCATG 480
Db 834 TCCAGTAATGACACCGTTCAACCTGACAGTATTAATGAAAGGCTTCTCACTCGGGCATG 893
Qy 481 ATCAACCCAGCGGCAAGTGGCAGACCTGACAGAACACGGCGGTGCGCACTTTGAG 540
Db 894 ATCAACCCAGCGGCAAGTGGCAGACCTGACAGAACACGGCGGTGCGCACTTTGAG 953
Qy 541 TATCAGATCCGCGTGACTGTGATGACTACTATATGGCTTTGGCTGYAATAAGTTCTGC 600
Db 954 TATCAGATCCGCGTGACTGTGATGACTACTATATGGCTTTGGCTGCAATAGTTCTGC 1013
Qy 601 CGCCCCAGAGATGACTTTTGGACACTATGCTCTGACAGAAATGGCAACAAACTTGC 660
Db 1014 CGCCCCAGAGATGACTTTTGGACACTATGCTCTGACAGAAATGGCAACAAACTTGC 1073
Qy 661 ATGGAGGCTGGATGGGCGCGAATGTAAACAGAGCTATTTGCGACAAAGGCTCAGTCT 720
Db 1074 ATGGAGGCTGGATGGGCGCGAATGTAAACAGAGCTATTTGCGACAAAGGCTCAGTCT 1133
Qy 721 AAGCATGGGTCTTGCAAACTCCAGGTGACTGACAGGTGCGAGTAYGGCTGGCAAGGCTG 780
Db 1134 AAGCATGGGTCTTGCAAACTCCAGGTGACTGACAGGTGCGAGTATGGCTGGCAAGGCTG 1193
Qy 781 TACTGTGATAGTGCATCCACACCCGGGATGCGTCCACGGCATCTGTAAATGAGCCCTGG 840
Db 1194 TACTGTGATAGTGCATCCACACCCGGGATGCGTCCACGGCATCTGTAAATGAGCCCTGG 1253
Qy 841 CAGTGCTCTGTGAGACCAACTGGGCGGCGAGCTCTGTGACAAAGATCTCAATTAAGTGT 900
Db 2334 GGTGGCATTGTCATCGATGGTGTCAACTCTCTCAAGTGCATCTGTAGTACGGCTGGAG 2393

Db 1254 CAGTGCTCTGTGAGACCAAACTGGGGCGCGAGCTCTGTGACAAAGATCTCAATTAAGTGT 1313
Qy 901 GGGACTCATCAGCGGTGTCTCAACGGGGGAACCTTGTAGCAACACAGCGCCCTGACAAATAT 960
Db 1314 GGGACTCATCAGCGGTGTCTCAACGGGGGAACCTTGTAGCAACACAGCGCCCTGACAAATAT 1373
Qy 961 CAGTGTCTGCTGCTGAGGGTATTCAGGACCCCACTGTGAAATTTGCTGAGCAGCCCTGC 1020
Db 1374 CAGTGTCTGCTGCTGAGGGTATTCAGGACCCCACTGTGAAATTTGCTGAGCAGCCCTGC 1433
Qy 1021 CTCTCTGATCCCTGTCAACACAGAGCAGCTGTAAAGAGACCTCCCTGGGCTTTGAGTGT 1080
Db 1434 CTCTCTGATCCCTGTCAACACAGAGCAGCTGTAAAGAGACCTCCCTGGGCTTTGAGTGT 1493
Qy 1081 GAGTGTTCCTGAGGCTGAGACCGGCGCCCACTGTCTCAAAACATTTGATGACTGTCTCCT 1140
Db 1494 GAGTGTTCCTGAGGCTGAGACCGGCGCCCACTGTCTCAAAACATTTGATGACTGTCTCCT 1553
Qy 1141 AATAACTGTTCCCAAGGGGCACTGTCAGGACCTGGTTAAACGGATTTAAGTGTGTGTC 1200
Db 1554 AATAACTGTTCCCAAGGGGCACTGTCAGGACCTGGTTAAACGGATTTAAGTGTGTGTC 1613
Qy 1201 CCCCCACAGTGGAGCTGGGAAAACGTCAGTGTAGTGCATAAATGAACTGTAGGGCCAAACCT 1260
Db 1614 CCCCCACAGTGGAGCTGGGAAAACGTCAGTGTAGTGCATAAATGAACTGTAGGGCCAAACCT 1673
Qy 1261 TGTGTAAACGCCAAATCTCTGTAAGAATCTCATTTGCCAGCTACTACTGCGAGTGTCTTCCC 1320
Db 1674 TGTGTAAACGCCAAATCTCTGTAAGAATCTCATTTGCCAGCTACTACTGCGAGTGTCTTCCC 1733
Qy 1321 GGTGATGAGGTGAGAAATTTGTGACATAAATATTAATGACTGCTTGGCCAGTGTCAAGAT 1380
Db 1734 GGTGATGAGGTGAGAAATTTGTGACATAAATATTAATGACTGCTTGGCCAGTGTCAAGAT 1793
Qy 1381 GAGCCTCTGCTGGGATTTGGTTAATGTTATCGCTGTATCTGTCCACCTGGCTATGCA 1440
Db 1794 GAGCCTCTGCTGGGATTTGGTTAATGTTATCGCTGTATCTGTCCACCTGGCTATGCA 1853
Qy 1441 GGGATCACTGTGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGGT 1500
Db 1854 GGGATCACTGTGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGGT 1913
Qy 1501 CACTGTCAAGATGAATCAACAGATTCAGTGTCTGTGTCCCACTGGTTTCTCTGGAAC 1560
Db 1914 CACTGTCAAGATGAATCAACAGATTCAGTGTCTGTGTCCCACTGGTTTCTCTGGAAC 1973
Qy 1561 CTCTGTCACTGACATCGATTTATGTAGCCTAAATCCCTGCGAGACCGGTGCCAGTGC 1620
Db 1974 CTCTGTCACTGACATCGATTTATGTAGCCTAAATCCCTGCGAGAACCGGTGCCAGTGC 2033
Qy 1621 TACAACCTGCCAGTGAATTTCTGCAAGTCCCGAGGACTATGAGGGCAAGAACTGC 1680
Db 2034 TACAACCTGCCAGTGAATTTCTGCAAGTCCCGAGGACTATGAGGGCAAGAACTGC 2093
Qy 1681 TACAACCTGAAAGACCACTGCGCGACGACCCCTGTGAAAGTGAATGACGTGCACAGTG 1740
Db 2094 TACAACCTGAAAGACCACTGCGCGACGACCCCTGTGAAAGTGAATGACGTGCACAGTG 2153
Qy 1741 GCCATGGCTTCCAAACGACACCTGAAGGGGTGGGTATATTTCTCCCAAGCTGTGGT 1800
Db 2154 GCCATGGCTTCCAAACGACACCTGAAGGGGTGGGTATATTTCTCCCAAGCTGTGGT 2213
Qy 1801 CCTCAGGGAAAGTGAAGAGTCACTGGGAGGCAAAATTCACCTGTGACTGTAAACAAAGGC 1860
Db 2214 CCTCAGGGAAAGTGAAGAGTCACTGGGAGGCAAAATTCACCTGTGACTGTAAACAAAGGC 2273
Qy 1861 TTCAACGGGAAACATCTGCGCATGAAATATTAATGACTGTGAGAGCAACCCCTGTGAAAC 1920
Db 2274 TTCAACGGGAAACATCTGCGCATGAAATATTAATGACTGTGAGAGCAACCCCTGTGAAAC 2333
Qy 1921 GGTGGCATTGCAATCGATGGTGTCAACTCTTACAAGTGCATCTGTAGTACGGCTGGAG 1980
Db 2334 GGTGGCATTGCAATCGATGGTGTCAACTCTTACAAGTGCATCTGTAGTACGGCTGGAG 2393

```
QY 1981 GGGSCCTACTGTGAACCAATATTTAATGACTGACGCCAGAACCCCTGCCACAATGGGGC 2040
Db 2394 GGGSCCTACTGTGAACCAATATTTAATGACTGACGCCAGAACCCCTGCCACAATGGGGC 2453
QY 2041 ACGTGTCGCGACCTGGTCAATGACTCTTACTGTGACTGTGTAAAATGGGTGGAAGAAAG 2100
Db 2454 ACGTGTCCGACCTGGTCAATGACTCTTACTGTGACTGTGTAAAATGGGTGGAAGAAAG 2513
QY 2101 ACTTGCCACTCACTGACAGTGTGATGAGGCCACGTGCAACAAACGGTGGCACTTGC 2160
Db 2514 ACTTGCCACTCACTGACAGTGTGATGAGGCCACGTGCAACAAACGGTGGCACTTGC 2573
QY 2161 TATGATAGGGGGATGCTTTAAGTGCATGTGTCTTGGCGCTGGGAAGCAACAACCTGT 2220
Db 2574 TATGATAGGGGGATGCTTTAAGTGCATGTGTCTTGGCGCTGGGAAGCAACAACCTGT 2633
QY 2221 AACATAGCCGAAACAGTAGCTGCTGCCCAACCCCTGCCATATATGGGGGCACATGTGTG 2280
Db 2634 AACATAGCCGAAACAGTAGCTGCTGCCCAACCCCTGCCATATATGGGGGCACATGTGTG 2693
QY 2281 GTCAACGGCGAGTCTTTAAGTGCATGTGTGCAAGGAAGGCTGGAGGGGCCCATCTGTGCT 2340
Db 2694 GTCAACGGCGAGTCTTTAAGTGCATGTGTGCAAGGAAGGCTGGAGGGGCCCATCTGTGCT 2753
QY 2341 CAGAAATACCAATGACTGCAGCCCTCATCCCTGTTTACAACAGCGGCACCTGTGTGGATGGA 2400
Db 2754 CAGAAATACCAATGACTGCAGCCCTCATCCCTGTTTACAACAGCGGCACCTGTGTGGATGGA 2813
QY 2401 GACAACTGGTACCCTGCGAATGTGCCCGGGTTTTGCTGGGCCCGAGCTCAGAAATAAC 2460
Db 2814 GACAACTGGTACCCTGCGAATGTGCCCGGGTTTTGCTGGGCCCGAGCTCAGAAATAAC 2873
QY 2461 ATCAATGAATGCCAGTCTTACCTTTGTCCTTTGGAGGACCTGTGTGGATGAGATCAAT 2520
Db 2874 ATCAATGAATGCCAGTCTTACCTTTGTCCTTTGGAGGACCTGTGTGGATGAGATCAAT 2933
QY 2521 GGCTACCGGTGTGCTGCCCTCCAGGGCACAGTGGTCCCAAGTGCACAGGAAGTTTTCAGGG 2580
Db 2934 GGCTACCGGTGTGCTGCCCTCCAGGGCACAGTGGTCCCAAGTGCACAGGAAGTTTTCAGGG 2993
QY 2581 AGACTCTTGATCACCATGGGAGTGTGATACAGATGGGGCCAAATGGGATGATGACTGT 2640
Db 2994 AGACTCTTGATCACCATGGGAGTGTGATACAGATGGGGCCAAATGGGATGATGACTGT 3053
QY 2641 AATACTGCCAGTGTGAATGGAACGGATCCCTGCTCAAGAGTCTGTGTGGCCCTCGA 2700
Db 3054 AATACTGCCAGTGTGAATGGAACGGATCCCTGCTCAAGAGTCTGTGTGGCCCTCGA 3113
QY 2701 CCTTGGCTGCTCCACAAAGGGCACAGGAGTGGCCCGGCGAGCTGCATCCCATC 2760
Db 3114 CCTTGGCTGCTCCACAAAGGGCACAGGAGTGGCCCGGCGAGCTGCATCCCATC 3173
QY 2761 CTGACGACCAAGTGTCTTCCCTCCAGCTGCTGTGGCGAGTGTGCTTCCAGT 2820
Db 3174 CTGACGACCAAGTGTCTTCCCTCCAGCTGCTGTGGCGAGTGTGCTTCCAGT 3233
QY 2821 CTCAGCCGGTGAAGACAAAGTGCATCTCTGATCTCTATTAACAGGATAAATCTGGCAAC 2880
Db 3234 CTCAGCCGGTGAAGACAAAGTGCATCTCTGATCTCTATTAACAGGATAAATCTGGCAAC 3293
QY 2881 ATCACAATTTACTTTTAAAGAGATGATGTCACCGGTCTTACTACGGACACATTTGC 2940
Db 3294 ATCACAATTTACTTTTAAAGAGATGATGTCACCGGTCTTACTACGGACACATTTGC 3353
QY 2941 AGTGAATTTAGGAATTTGAATATTTTGAAGAAATGTTTCCGTGAAATATTTCAATCTACATC 3000
Db 3354 AGTGAATTTAGGAATTTGAATATTTTGAAGAAATGTTTCCGTGAAATATTTCAATCTACATC 3413
QY 3001 GCTTGGAGCCTTCCCTTTCAGCGCAACAATGAATAATATGATGGCCATTTCTGTGTAAGAT 3060
Db 3414 GCTTGGAGCCTTCCCTTTCAGCGCAACAATGAATAATATGATGGCCATTTCTGTGTAAGAT 3473
```

```
QY 3061 ATACGGGATGATGGGAACCCGATCAAGAAATCACTGACAAATAATATCGATCTTGTAGT 3120
Db 3474 ATACGGGATGATGGGAACCCGATCAAGAAATCACTGACAAATAATATCGATCTTGTAGT 3533
QY 3121 AAACGTGATGGAACACAGCTCGCTGATTTGCTGCGTTGCAAGTAAGAGTTCAGAGCGG 3180
Db 3534 AAACGTGATGGAACACAGCTCGCTGATTTGCTGCGTTGCAAGTAAGAGTTCAGAGCGG 3593
QY 3181 CCTCTGAAGAACAGAACAGATTTCCCTTGTTCCTTCTGAGCTCTCTCTTAACTGTGGCT 3240
Db 3594 CCTCTGAAGAACAGAACAGATTTCCCTTGTTCCTTCTGAGCTCTCTCTTAACTGTGGCT 3653
QY 3241 TGGATCTGTGCTTGGTGAACGCTTCTTACTTGTGCTGCGGAAGCGGGAAGCCGGC 3300
Db 3654 TGGATCTGTGCTTGGTGAACGCTTCTTACTTGTGCTGCGGAAGCGGGAAGCCGGC 3713
QY 3301 AGCCACACACACTCAGCGCTCTGAGGACAAACACCAACAAACGTGGGGAGAGCTGAAC 3360
Db 3714 AGCCACACACACTCAGCGCTCTGAGGACAAACACCAACAAACGTGGGGAGAGCTGAAC 3773
QY 3361 CAGATCAAAAACCCCAATTTGAGAAACATGGGCCAAACACCGTCCCATCAAGGATTACGAG 3420
Db 3774 CAGATCAAAAACCCCAATTTGAGAAACATGGGCCAAACACCGTCCCATCAAGGATTACGAG 3833
QY 3421 AACAAAGAACTCCAAAATGTCTAAAATTAAGGACACACAATTTCTGAAGTAGAAGAGGACGAC 3480
Db 3834 AACAAAGAACTCCAAAATGTCTAAAATTAAGGACACACAATTTCTGAAGTAGAAGAGGACGAC 3893
QY 3481 ATGGACAAACACAGACAGAAAGCCCGTTTGGCAAGCAGCGCGGTATAGCTGTAGAC 3540
Db 3894 ATGGACAAACACAGACAGAAAGCCCGTTTGGCAAGCAGCGCGGTATAGCTGTGTAGAC 3953
QY 3541 AGAAGAGAGAGAGCCCGCCCAACGCGACGCGACAAAAACACCAAACTGGACAAAACAAACAG 3600
Db 3954 AGAAGAGAGAGAGCCCGCCCAACGCGACGCGACAAAAACACCAAACTGGACAAAACAAACAG 4013
QY 3601 GACACAGAGACTTGGAAAGTGCACAGAGCTTAAACGAATGAGGTACATCGTATAG 3657
Db 4014 GACACAGAGACTTGGAAAGTGCACAGAGCTTAAACGAATGAGGTACATCGTATAG 4070
```

```
RESULT 11
AR182312
LOCUS AR182312 4208 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 10 from patent US 6337387.
ACCESSION AR182312
VERSION AR182312.1 GI:20225228
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4208)
AUTHORS Sakano,S. and Itoh,A.
TITLE Differentiation-suppressive polypeptide
JOURNAL Patent: US 6337387-A 10 08-JAN-2002;
FEATURES Location/Qualifiers
source 1..4208
/mol_type="unassigned DNA"
```

```
ORIGIN
Query Match 99.9%; Score 3651; DB 6; Length 4208;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3651; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 ATGCGTTTCCCAACGACACRCCGCGCGTCCGGGCGCCCTTAAGCCTTCCTGCTCGCCCTG 60
Db 409 ATGCGTTTCCCAACGACACRCCGCGCGTCCGGGCGCCCTTAAGCCTTCCTGCTCGCCCTG 468
QY 61 CTCTGTGCTTGGAGCAAGGTGTGTGGGCGCTCGGGCTCAGTTTCGAGTTGGAGATCTG 120
Db 469 CTCTGTGCTTGGAGCAAGGTGTGTGGGCGCTCGGGCTCAGTTTCGAGTTGGAGATCTG 528
```

QY 121 TCCATGACGAA CGTGAA CGGGGAGCTGCAGAA CGGGAAC TGCTGCGGCGGCGCCGGAAC 180
DB 529 TCCATGACGAA CGTGAA CGGGGAGCTGCAGAA CGGGAAC TGCTGCGGCGGCGCCGGAAC 588
QY 181 CCGGAGACCGAAGTGCACCCCGGAGAGTGAGACATACATCTTCAAAGTGTCCTCAAG 240
DB 589 CCGGAGACCGAAGTGCACCCCGGAGAGTGAGACATACATCTTCAAAGTGTCCTCAAG 648
QY 241 GAGTATCAGTCCGCGTCAAGCGCGGGGGCCCTGCGAGCTTCGGCTCAGGGTCCAGCCCT 300
DB 649 GAGTATCAGTCCGCGTCAAGCGCGGGGGCCCTGCGAGCTTCGGCTCAGGGTCCAGCCCT 708
QY 301 GTCATCGGGGCAACACCTTCAACCTCAAGCCGAGCGCGGCAACGACCCCAACCGCATC 360
DB 709 GTCATCGGGGCAACACCTTCAACCTCAAGCCGAGCGCGGCAACGACCCCAACCGCATC 768
QY 361 GTGCTGCTTTTCAAGTTTCGCTGCGCGAGGTCTCTATACGTTTGTGAGGGGTGGAT 420
DB 769 GTGCTGCTTTTCAAGTTTCGCTGCGCGAGGTCTCTATACGTTTGTGAGGGGTGGAT 828
QY 421 TCCAGTAATGACACCGTTCAACCTGACAGTATTAATTGAAAGGCTTCTCACTCGGGCATG 480
DB 829 TCCAGTAATGACACCGTTCAACCTGACAGTATTAATTGAAAGGCTTCTCACTCGGGCATG 888
QY 481 ATCAACCCGAGCGGAGTGCGAGACGCTGAGCAGAAACCGGGCGTTGCCACTTTGAG 540
DB 889 ATCAACCCGAGCGGAGTGCGAGACGCTGAGCAGAAACCGGGCGTTGCCACTTTGAG 948
QY 541 TATCAGATCCGCTGACCTGTGATGACTACTATAGCTTTTGGCTTGGCTGTAATAGTTCTGC 600
DB 949 TATCAGATCCGCTGACCTGTGATGACTACTATAGCTTTTGGCTTGGCTGTAATAGTTCTGC 1008
QY 601 CGCCCCAGAGATGACTTTCTTTGGACACTATGCTCTGTGACAGAAATGGCAACAAACTTGC 660
DB 1009 CGCCCCAGAGATGACTTTCTTTGGACACTATGCTCTGTGACAGAAATGGCAACAAACTTGC 1068
QY 661 ATGGAAGGCTGGATGGGCCCGCAATGTAAACAGAGCTATTTTGGCAACAGGCTGAGTCT 720
DB 1069 ATGGAAGGCTGGATGGGCCCGCAATGTAAACAGAGCTATTTTGGCAACAGGCTGAGTCT 1128
QY 721 AAGCATGGGTCTTGCAAACTCCAGGTGACTGAGGTGCGAGTAYGGCTGGCAGGCTG 780
DB 1129 AAGCATGGGTCTTGCAAACTCCAGGTGACTGAGGTGCGAGTAYGGCTGGCAGGCTG 1188
QY 781 TACTGTGATAAGTGATCCACACCCGGGATGCTCCACGGCATCTGTAATGAGCCCTGG 840
DB 1189 TACTGTGATAAGTGATCCACACCCGGGATGCTCCACGGCATCTGTAATGAGCCCTGG 1248
QY 841 CAGTGCTCTGTGAGACCAACTGGGGCGGCGAGCTCTGTGACAAAGATCTCAATTAAGT 900
DB 1249 CAGTGCTCTGTGAGACCAACTGGGGCGGCGAGCTCTGTGACAAAGATCTCAATTAAGT 1308
QY 901 GGGACTCATCAGCGGTGTCTCAACGGGGGAACTTGTAGCAACACAGGCGCTGCAATAT 960
DB 1309 GGGACTCATCAGCGGTGTCTCAACGGGGGAACTTGTAGCAACACAGGCGCTGCAATAT 1368
QY 961 CAGTGTCTCCCTGAGGGGTATTTACAGACCCAACTGTGCAAAATTTGTGAGCAGCGCTGC 1020
DB 1369 CAGTGTCTCCCTGAGGGGTATTTACAGACCCAACTGTGCAAAATTTGTGAGCAGCGCTGC 1428
QY 1021 CTCTGTGATCCCTGTCAACAGAGGCGAGCTGTAAGGAGACCTCCCTGGGCTTTGAGTGT 1080
DB 1429 CTCTGTGATCCCTGTCAACAGAGGCGAGCTGTAAGGAGACCTCCCTGGGCTTTGAGTGT 1488
QY 1081 GAGTGTTCCTCAGGCTGGACCGGGCCGACATGCTCTCAAAACATTTGATGATGCTTCTCCT 1140
DB 1489 GAGTGTTCCTCAGGCTGGACCGGGCCGACATGCTCTCAAAACATTTGATGATGCTTCTCCT 1548
QY 1141 AATAACTGTTTCCACAGGGGCACTGCGAGGACTGTTTAAACGATTTAAAGTGTGTGTC 1200
DB 1549 AATAACTGTTTCCACAGGGGCACTGCGAGGACTGTTTAAACGATTTAAAGTGTGTGTC 1608
QY 1201 CCCCCACAGTGGAGTGGGAAAAACGTGCCAGTTAGATGCAAAATGAATGTGAGGGCCAAACCT 1260

DB 1609 CCCCCACAGTGGAGTGGGAAAAACGTGCCAGTTAGATGCAAAATGAATGTGAGGGCCAAACCT 1668
QY 1261 TGTGTAAAACGCCAAATCTGTGAAGAACTCTCATTTGCCAGCTACTACTCGCAGCTGTCTTCCC 1320
DB 1669 TGTGTAAAACGCCAAATCTGTGAAGAACTCTCATTTGCCAGCTACTACTCGCAGCTGTCTTCCC 1728
QY 1321 GGCTGAGTGGGTGAGAAATTTGTGATGATAAATATTAATTAATGATGCTTGGCCAGTGTGAGAA 1380
DB 1729 GGCTGAGTGGGTGAGAAATTTGTGATGATAAATATTAATTAATGATGCTTGGCCAGTGTGAGAA 1788
QY 1381 GAGCCTCTCTGTCGGGATTTGGTTAATGTTATCGCTGTATCTGTCCACCTGGCTATGCA 1440
DB 1789 GAGCCTCTCTGTCGGGATTTGGTTAATGTTATCGCTGTATCTGTCCACCTGGCTATGCA 1848
QY 1441 GGCATCACTGTGAGAGACATCGATGAATGTGCCAGCAACCCCTGTGTTGAATGGGGGT 1500
DB 1849 GGCATCACTGTGAGAGACATCGATGAATGTGCCAGCAACCCCTGTGTTGAATGGGGGT 1908
QY 1501 CACTGTGAGAAATCAACAGATTCAGATTCAGTGTCTGTGCCCACTGGTTTCTCTGGAAC 1560
DB 1909 CACTGTGAGAAATCAACAGATTCAGATTCAGTGTCTGTGCCCACTGGTTTCTCTGGAAC 1968
QY 1561 CTCTGTGAGTGCACATCGATTTATTTGTGAGCTTAATCCCTGCGCAGAACCGTGCCAGTGC 1620
DB 1969 CTCTGTGAGTGCACATCGATTTATTTGTGAGCTTAATCCCTGCGCAGAACCGTGCCAGTGC 2028
QY 1621 TACAAACCGTGCAGTGTGACTATTTCTGCAAGTGCCCGAGGACTATGAGGGCAAGAACTGC 1680
DB 2029 TACAAACCGTGCAGTGTGACTATTTCTGCAAGTGCCCGAGGACTATGAGGGCAAGAACTGC 2088
QY 1681 TCACACTGAAAGACCACTGCGCAGCAACCCCTGTGAAGTGTATGACAGCTGCACAGTG 1740
DB 2089 TCACACTGAAAGACCACTGCGCAGCAACCCCTGTGAAGTGTATGACAGCTGCACAGTG 2148
QY 1741 GGCATGGCTTCCAAACGACACACCTGAAGGGGTGGGTATTTCTTCCCAAGCTGTGGT 1800
DB 2149 GGCATGGCTTCCAAACGACACACCTGAAGGGGTGGGTATTTCTTCCCAAGCTGTGGT 2208
QY 1801 CCTCAGGGAAGTGCAGAGTCACTCGGGAGGCAAAATTCACCTGTGACTGTAAACAAAGGC 1860
DB 2209 CCTCAGGGAAGTGCAGAGTCACTCGGGAGGCAAAATTCACCTGTGACTGTAAACAAAGGC 2268
QY 1861 TTCACGGGAAACATPACTGCGCAATGAAAATATTAATGACTGTGAGAGCAACCCCTGTGAGAAC 1920
DB 2269 TTCACGGGAAACATPACTGCGCAATGAAAATATTAATGACTGTGAGAGCAACCCCTGTGAGAAC 2328
QY 1921 GGTGGCACTTGCATCGATGGTGTCAACTCTCAAGTGCATCTGTAGTGAAGCTGGAG 1980
DB 2329 GGTGGCACTTGCATCGATGGTGTCAACTCTCAAGTGCATCTGTAGTGAAGCTGGAG 2388
QY 1981 GGGGCTTACTGTGAAACCAATATTAATGACTGAGCAGCAACCCCTGCCACAATGGGGGC 2040
DB 2389 GGGGCTTACTGTGAAACCAATATTAATGACTGAGCAGCAACCCCTGCCACAATGGGGGC 2448
QY 2041 ACCTGTGCGGACCTGGTCAATGACTTCTACTGTGACTGTAAAAATGGGTGAAAAGGAAAG 2100
DB 2449 ACCTGTGCGGACCTGGTCAATGACTTCTACTGTGACTGTAAAAATGGGTGAAAAGGAAAG 2508
QY 2101 ACCTGCACTCAGCTGACAGTGTGATGAGGCCACGTCGCAACCAACCGTGGCACCTGC 2160
DB 2509 ACCTGCACTCAGCTGACAGTGTGATGAGGCCACGTCGCAACCAACCGTGGCACCTGC 2568
QY 2161 TATGATGAGGGGATCTTTAAGTGCATGTGTCTTGGGGCTGGGAAGGAAACCAACCTGT 2220
DB 2569 TATGATGAGGGGATCTTTAAGTGCATGTGTCTTGGGGCTGGGAAGGAAACCAACCTGT 2628
QY 2221 AACATAGCCGAAACAGTGTGCTGCCCAACCCCTGCCCAATATGGGGGCAATGTGTG 2280
DB 2629 AACATAGCCGAAACAGTGTGCTGCCCAACCCCTGCCCAATATGGGGGCAATGTGTG 2688
QY 2281 GTCAAAGGCGAGTCTTTTACGTGCTGTGCAAGGAAGCTGGGAGGGGCCCATCTGTGCT 2340

Db 2689 GTCAACGGCGAGTCTTTTACGTGCGTCTGCAAGAAAGCGTGGAGGGGCCCATCTGTGCT 2748
Qy 2341 CAGAAATACCAATGATCGACGCCCTCATCCCTGTGTTAACAAGCGGCACCTGTGTGGATGGA 2400
Db 2749 CAGAATACCAATGATCGACGCCCTCATCCCTGTGTTAACAAGCGGCACCTGTGTGGATGGA 2808
Qy 2401 GACAACCTGGTACGGTGCAGAAATGTGCCCGGGTTTTGTGCGGCCGAGCTCAGAAATAAAC 2460
Db 2809 GACAACCTGGTACGGTGCAGAAATGTGCCCGGGTTTTGTGCGGCCGAGCTCAGAAATAAAC 2868
Qy 2461 ATCAATGAATGCCAGTCTTACACCTGTGTCCTTTGGAGCGACCTGTGTGGATGATCAAT 2520
Db 2869 ATCAATGAATGCCAGTCTTACACCTGTGTCCTTTGGAGCGACCTGTGTGGATGATCAAT 2928
Qy 2521 GGCTACCGGTGTGTCCTGCTCCAGGGCACAAGTGGTGCAGTGCAGAGAGTTTCAGGG 2580
Db 2929 GGCTACCGGTGTGTCCTGCTCCAGGGCACAAGTGGTGCAGTGCAGAGAGTTTCAGGG 2988
Qy 2581 AGACCTTGCATCACATGGGAGTGTGATACAGATGGGCGCCAAATGGGATGATGACTGT 2640
Db 2989 AGACCTTGCATCACATGGGAGTGTGATACAGATGGGCGCCAAATGGGATGATGACTGT 3048
Qy 2641 AATACCTGCCAGTGCCTGAATGGACCGATCGCTGCTCAAAGGTCTGGTGTGGCCCTCGA 2700
Db 3049 AATACCTGCCAGTGCCTGAATGGACCGATCGCTGCTCAAAGGTCTGGTGTGGCCCTCGA 3108
Qy 2701 CCTTGCTGCTCCAAAGGGCAGCGAGTGCCTGAGCGGGCAGAGCTGCATCCCATC 2760
Db 3109 CCTTGCTGCTCCAAAGGGCAGCGAGTGCCTGAGCGGGCAGAGCTGCATCCCATC 3168
Qy 2761 CTGAGCAGCAGTGCCTTCTGCTCACCTGCACTGGTGTGGCGAGTGTGGTCTTCAGT 2820
Db 3169 CTGAGCAGCAGTGCCTTCTGCTCACCTGCACTGGTGTGGCGAGTGTGGTCTTCAGT 3228
Qy 2821 CTCCAGCGGTGAAGACAAAGTGCACCTCTGACTCTTATACCAAGGATAACTGTGCGAAC 2880
Db 3229 CTCCAGCGGTGAAGACAAAGTGCACCTCTGACTCTTATACCAAGGATAACTGTGCGAAC 3288
Qy 2881 ATCAATTTACCTTAAACAGGAGATGATGCACCGTCTTACTACGAGACATTTGC 2940
Db 3289 ATCAATTTACCTTAAACAGGAGATGATGCACCGTCTTACTACGAGACATTTGC 3348
Qy 2941 AGTGAATGGAGAAATTTGAATATTTGAAGAAATGTTTCGCTGAAATTAATCAATCATC 3000
Db 3349 AGTGAATGGAGAAATTTGAATATTTGAAGAAATGTTTCGCTGAAATTAATCAATCATC 3408
Qy 3001 GCTTGCAGCGCTTCCCTTACAGGACAAATGAATAATATGCGCAATTTCTGCTGAAGAT 3060
Db 3409 GCTTGCAGCGCTTCCCTTACAGGACAAATGAATAATATGCGCAATTTCTGCTGAAGAT 3468
Qy 3061 ATACGGGATGATGGAAACCGGATCAAGGAATCACTGCAAAATTAATCGATCTGTGTAGT 3120
Db 3469 ATACGGGATGATGGAAACCGGATCAAGGAATCACTGCAAAATTAATCGATCTGTGTAGT 3528
Qy 3121 AAAAGTGATGAAACAGCTCCCTGATGCTCCCTGTGAGAAAGTTCAGAGCGG 3180
Db 3529 AAAAGTGATGAAACAGCTCCCTGATGCTCCCTGTGAGAAAGTTCAGAGCGG 3588
Qy 3181 CCTCTGAAGAAACAGAGATTTTCTGTGTTCCCTTGTGAGCTCTGTCTTAATCTGTGCT 3240
Db 3589 CCTCTGAAGAAACAGAGATTTTCTGTGTTCCCTTGTGAGCTCTGTCTTAATCTGTGCT 3648
Qy 3241 TGGATCTGTTGTTGGTGAAGGCTTCTACTGTTGCTGCGGAAAGCGGCGGAAACCGGGC 3300
Db 3649 TGGATCTGTTGTTGGTGAAGGCTTCTACTGTTGCTGCGGAAAGCGGCGGAAACCGGGC 3708
Qy 3301 AGCCACACACTCAGCTCTTGAAGACAAACACCAACAGCTGCGGAGCAGCTGAAC 3360
Db 3709 AGCCACACACTCAGCTCTTGAAGACAAACACCAACAGCTGCGGAGCAGCTGAAC 3768
Qy 3361 CAGATCAAAAAACCCATTGAGAAACATGGGCGCAACAGCGTCCCATCAAGGATTACGAG 3420
Db 3769 CAGATCAAAAAACCCATTGAGAAACATGGGCGCAACAGCGTCCCATCAAGGATTACGAG 3828

Qy 3421 AACAGAACTCCAAATGTCTAAATTAAGGACACACAATTTCTGAAGTAGAAGAGGACGAC 3480
Db 3829 AACAGAACTCCAAATGTCTAAATTAAGGACACACAATTTCTGAAGTAGAAGAGGACGAC 3888
Qy 3481 ATGGACAAACACACAGCAGAAAGCCCGTTTGGCAAGCAGCCGCGTATACGCTGGTAGAC 3540
Db 3889 ATGGACAAACACACAGCAGAAAGCCCGTTTGGCAAGCAGCCGCGTATACGCTGGTAGAC 3948
Qy 3541 AGAAGAAAGAGCGCCCGCCCAACGGCAGCGCAGCAAAACACCCAACTTGGACAAACAAACAG 3600
Db 3949 AGAAGAAAGAGCGCCCGCCCAACGGCAGCGCAGCAAAACACCCAACTTGGACAAACAAACAG 4008
Qy 3601 GACAAACAGAGACTTGGAAAGTGCCAGAGCTTAAACCGGAATGGAGTACATCGTATAG 3657
Db 4009 GACAAACAGAGACTTGGAAAGTGCCAGAGCTTAAACCGGAATGGAGTACATCGTATAG 4065

RESULT 12
AF028593
LOCUS
DEFINITION
Homo sapiens transmembrane protein Jagged 1 (HJ1) mRNA, complete cds.
ACCESSION
VERSION
AF028593
AF028593.1 GI:2599081
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 5457)
Lindsell,C.E., Shawber,C.J., Boulter,J. and Weinmaster,G.
Jagged: a mammalian ligand that activates Notch1
Cell 80 (6), 909-917 (1995)
PUBMED
7697721
REFERENCE
2 (bases 1 to 5457)
Bash,J., Zong,W.X., Banga,S., Rivera,A., Ballard,D.W., Ron,Y. and Gelinias,C.
Rel/NF-kappaB can trigger the Notch signaling pathway by inducing the expression of Jagged1, a ligand for Notch receptors
EMBO J. 18 (10), 2803-2811 (1999)
PUBMED
10329626
REFERENCE
3 (bases 1 to 5457)
Bash,J., Zong,W.-X. and Gelinias,C.
Direct Submission
Submitted (03-OCT-1997) Center for Advanced Biotechnology and Medicine, Department of Biochemistry, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, 679 Hoes Lane, Piscataway, NJ 08854, USA

FEATURES
source

1..5457
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20p12"
/cell_line="HeLa cells"
1..5457
/gene="HJ1"
/genes="HJ1"
/notes="similar to H. sapiens Jagged 1 encoded by GenBank Accession Numbers U73936, AF003837, and U77720"
/codon_start=1
/product="transmembrane protein Jagged 1"
/protein_id="AAB84053.1"
/db_xref="GI:2599082"
/translation="MRSPTRGSRPLSLALLCALRAKVCAGSQFLEILSMON VNGELQNGCCGARNPDKCTRSDCDYFKVCLKEYQSRVTAGGPCSFGSGSTPVI GGNFTNLKASGRNDRNRIVLFPFANPRSYLLVEANDSDNDTVQPDLSIEKASHSGM INFSRQWTLKQNTGVAFYQIRVTCDDYDYGKNCFKRPRDFFGHATCDONGNK TCMEGWMGPENRAICRQGSFKHSGCKLPDRCQYGWQGLYKCDKICIPHPGCVHIGC

gene

CDS

NEPQCLCETNHWGQLCDKDLNYCGTHQPCLINGTCSNTGPKDKVQCSCEBYSQPNCE
IAEHASLDFCHNRGSKETSLSGFECECSFGWTGPTCTSNIDDCSNPNCHSGTCCQDL
VNGFKVCPQPTWTKTQCLDANECEKPCVNAKSKNLNIAASYIDCCLPMMWQNCIDIN
INDLGGQCNDAKRLDINGYRCIPPGYAGDHCERDIDEASNPCLNGCHGCKNEINR
FOCLCTGFSGNLCQIDIDYCEPNPCONGAQCYNRASDYFKCPEDYEGKNCSEHLKH
CRTPPEVIDCTVAMASNDTPBQVYIISNVCGPHGKCKSQSGKFTCDCKNGFTGT
YCHENINDCESNCRNGGTCIDGVNSYKICSDGMEGAYCETNINDCSQNPCHNGGTC
RDIVNPFYCDCKNGWKTKCHSDSDQDEATCNNGGTCYDEGDAFKMCPGMEGTC
NIARNSCLPNCHNGTCVNVGESFTVCCKEWEGPICAOQNTNDCSPHPYNSNGTCV
DGNWYRCECAFAGPDCRININEQSSPCAFGATCVDREINGYRCVCPGPHSGAKCO
EVSGRCPITMGSVIPDCAKWDCCDTCOCLNGRIAGSKVMCGRPCLLKHKGSECPSSG
QSCIPILDDQCFVHPCTGCVGCHSSLOPVKTKTSDSYODNCANITFTFNKEMSP
GLTTEHICSELRNLILKNVSAEYSIYIACEPSPSANNEIHVAISREDIRDDGNPIKE
ITDKIIDLKSRDGNSSLIAAVAEVRVQRRLKRTDFLVPFLUSSVTPIMVYICCLVTA
FYWKLRRKPGSHTASIEDNTNNVRLQNIKNPIEKHGANTVPIIKDYENKNSKM
SKIRTNSEVEEDMDKHQOKARFAKQPAYTDLVREBKPNGTPTKHPNWNKQDNRD
LESAQSLNRMEYIV"

ORIGIN

Query Match 99.9%; Score 3651; DB 8; Length 5457;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3651; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy	1	ATGCGTTCCCGACGACRCGCGCGGTCCGGCGCGCCCTAAGCCTCTGCTCGCCCTG 60
Db	264	ATGCGTTCCCGACGACRCGCGCGGTCCGGCGCGCCCTAAGCCTCTGCTCGCCCTG 323
Qy	61	CTCTGTGCTCCGAGCAAGGTGTGTGGGCGCTCGGGTCAGTTCGAGTTGGAGATCCTG 120
Db	324	CTCTGTGCTCCGAGCAAGGTGTGTGGGCGCTCGGGTCAGTTCGAGTTGGAGATCCTG 383
Qy	121	TCCATGACAGAACGTGAACGGGAGCTGCAGAACGGGAACTGTCTGGCGGCGCGCCGGAAC 180
Db	384	TCCATGACAGAACGTGAACGGGAGCTGCAGAACGGGAACTGTCTGGCGGCGCGCCGGAAC 443
Qy	181	CCGGAGACCGAGGTGACCCCGCAGAGTGTGACATACATCTCAAAGTGTGCTCAAG 240
Db	444	CCGGAGACCGAGGTGACCCCGCAGAGTGTGACATACATCTCAAAGTGTGCTCAAG 503
Qy	241	GAGTATCAGTCCGCGCTCAAGCGCGGGGCGCTTGACGCTTCGGCTCAGGGTCCAGGCCT 300
Db	504	GAGTATCAGTCCGCGCTCAAGCGCGGGGCGCTTGACGCTTCGGCTCAGGGTCCAGGCCT 563
Qy	301	GTCAATCGGGGGAACACCTTCAACCTCAAGGCGAGCGCGGCAACGACCGCAACCGGATC 360
Db	564	GTCAATCGGGGGAACACCTTCAACCTCAAGGCGAGCGCGGCAACGACCGCAACCGGATC 623
Qy	361	GTGCTGCTTTTCAAGTTTCGCTCGCGAGGTCTTATACGTTGCTTGTGGAGGCGTGGAT 420
Db	624	GTGCTGCTTTTCAAGTTTCGCTCGCGAGGTCTTATACGTTGCTTGTGGAGGCGTGGAT 683
Qy	421	TCCAGTAATGACACCGTTCAACCTGACAGTATTATTGAAGGCTTCTCACTCGGCGATG 480
Db	684	TCCAGTAATGACACCGTTCAACCTGACAGTATTATTGAAGGCTTCTCACTCGGCGATG 743
Qy	481	ATCAACCCGAGCGGAGTGGAGACGCTGAAGCAGAACACGGCGGTGCGCACTTTGAG 540
Db	744	ATCAACCCGAGCGGAGTGGAGACGCTGAAGCAGAACACGGCGGTGCGCACTTTGAG 803
Qy	541	TATCAGATCCGCGTGACCTGTGATGACTACTATGCTGCTTGGCTGYAATAGTTCTGC 600
Db	804	TATCAGATCCGCGTGACCTGTGATGACTACTATGCTGCTTGGCTGYAATAGTTCTGC 863
Qy	601	CGCCCCGAGAGTACATTTTGGACACTATGCTCTGACAGAAATGGCAACAAACTTGC 660
Db	864	CGCCCCGAGAGTACATTTTGGACACTATGCTCTGACAGAAATGGCAACAAACTTGC 923
Qy	661	ATGGAAGGCTGAGTGGGCCCCGAATGTAAACAGAGCTATTTCGCAACAGGCTCAGTCTT 720
Db	924	ATGGAAGGCTGAGTGGGCCCCGAATGTAAACAGAGCTATTTCGCAACAGGCTCAGTCTT 983
Qy	721	AGCATGGGCTTGTGCAATCTCCAGGTGACTGACGCTGCGAGTAYGGCTGGCAAGGCTG 780

Db	984	AAGCATGGGCTTGTGCAAACTCCACAGGTGACTGCAAGGTGCCAGTACGGCTGCCAAGGCGCTG 1043
Qy	781	TACTGTGATAAGTGATCTCCACACCCGGGATGGGTCCACGGGCACTCTGTAAATGAGCCCTGG 840
Db	1044	TACTGTGATAAGTGATCTCCACACCCGGGATGGGTCCACGGGCACTCTGTAAATGAGCCCTGG 1103
Qy	841	CAGTGCTCTGTGAGACCAACTGGGGGGGCGACCTCTGTGACAAAGATCTCAATTACTGT 900
Db	1104	CAGTGCTCTGTGAGACCAACTGGGGGGGCGACCTCTGTGACAAAGATCTCAATTACTGT 1163
Qy	901	GGGACTCATCAGCGCTGTCTCAACGGGGGAATCTTGTAGCAACACAGGCCCTCGACAAATAT 960
Db	1164	GGGACTCATCAGCGCTGTCTCAACGGGGGAATCTTGTAGCAACACAGGCCCTCGACAAATAT 1223
Qy	961	CAGTGTTCCTGCCCTGAGGGGTAATTCAGGACCCAACTGTGAAATTCGTGAGCAAGCGCTGC 1020
Db	1224	CAGTGTTCCTGCCCTGAGGGGTAATTCAGGACCCAACTGTGAAATTCGTGAGCAAGCGCTGC 1283
Qy	1021	CTCTCTGATCCCTGTGTCACACAGAGGAGCTGTAAAGAGACCTCCCTGGGCTTTGAGTGT 1080
Db	1284	CTCTCTGATCCCTGTGTCACACAGAGGAGCTGTAAAGAGACCTCCCTGGGCTTTGAGTGT 1343
Qy	1081	GAGTGTTCCTCCAGGCTGGACCGGCCCCACATGCTCTACAAACATTTGATGACTGTTCCT 1140
Db	1344	GAGTGTTCCTCCAGGCTGGACCGGCCCCACATGCTCTACAAACATTTGATGACTGTTCCT 1403
Qy	1141	AATAACTGTTCCTCCAGGCGACCTGCCAGGACCTGTGTTAAAGGATTTAAAGTGTGTGC 1200
Db	1404	AATAACTGTTCCTCCAGGCGACCTGCCAGGACCTGTGTTAAAGGATTTAAAGTGTGTGC 1463
Qy	1201	CCCCCAGTGGACTGGGAAACGTCAGTTAGATGCAAAATGTAATGTTGAGGGCAACCT 1260
Db	1464	CCCCCAGTGGACTGGGAAACGTCAGTTAGATGCAAAATGTAATGTTGAGGGCAACCT 1523
Qy	1261	TGTGTAACCGCAAACTCTGTAAGATCTCATTTGCCAGCTACTACTGCGAGCTGTCTTCCC 1320
Db	1524	TGTGTAACCGCAAACTCTGTAAGATCTCATTTGCCAGCTACTACTGCGAGCTGTCTTCCC 1583
Qy	1321	GGCTGATGGGTGAGAAATTTGACATAAATATTAATGACTGCGCTTGGCCAGTGTGCAAT 1380
Db	1584	GGCTGATGGGTGAGAAATTTGACATAAATATTAATGACTGCGCTTGGCCAGTGTGCAAT 1643
Qy	1381	GAGCCTCTCTGCGGATTTGGTTAATGGTTATCGCTGTATCTGTGCACTGTGCTATGCA 1440
Db	1644	GAGCCTCTCTGCGGATTTGGTTAATGGTTATCGCTGTATCTGTGCACTGTGCTATGCA 1703
Qy	1441	GGCGATCATCTGTGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGT 1500
Db	1704	GGCGATCATCTGTGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGT 1763
Qy	1501	CACGTGTGAGAAATCAACAGATTCACAGTGTCTGTGTGCCACTGGTTTCTCTGGAAC 1560
Db	1764	CACGTGTGAGAAATCAACAGATTCACAGTGTCTGTGTGCCACTGGTTTCTCTGGAAC 1823
Qy	1561	CTCTGTGAGTGAATCGAATTTGTGAGCTTAATCCCTGCCAGAACCGTGGCCAGTGC 1620
Db	1824	CTCTGTGAGTGAATCGAATTTGTGAGCTTAATCCCTGCCAGAACCGTGGCCAGTGC 1883
Qy	1621	TACAAACCGTGCAGTGAATTTCTGCAAGTGCCTGAGGAGCTATGAGGGCAAGAACTGC 1680
Db	1884	TACAAACCGTGCAGTGAATTTCTGCAAGTGCCTGAGGAGCTATGAGGGCAAGAACTGC 1943
Qy	1681	TCACACTGAAAGACCACTGCGCGACGACCCCTGTGAAAGTGAATTCAGAGTGCACAGTG 1740
Db	1944	TCACACTGAAAGACCACTGCGCGACGACCCCTGTGAAAGTGAATTCAGAGTGCACAGTG 2003
Qy	1741	GCCATGGCTTCCAAACGACACCTGAAGGGGTGCGGTATATTTCTCCAAAGTGTGTGGT 1800
Db	2004	GCCATGGCTTCCAAACGACACCTGAAGGGGTGCGGTATATTTCTCCAAAGTGTGTGGT 2063
Qy	1801	CCTCAGGGAAGTGAAGAGTCACTGGGAGGCAAAATTCACCTGTGCTTACCAAGGC 1860
Db	2064	CCTCAGGGAAGTGAAGAGTCACTGGGAGGCAAAATTCACCTGTGCTTACCAAGGC 2123

QY	1861	TTCACGGAACATACCTGCGCATGAATAATTAATGACTGTGAGAGCAACCCCTTGTAGAAC	1920
Db	2124	TTCACGGAACATACCTGCGCATGAATAATTAATGACTGTGAGAGCAACCCCTTGTAGAAC	2183
QY	1921	GGTGGCACTTGCATCGATGGTGTCAACTCCCTACAAGTGCATCTGTAGTACGCGCTGGAG	1980
Db	2184	GGTGGCACTTGCATCGATGGTGTCAACTCCCTACAAGTGCATCTGTAGTACGCGCTGGAG	2243
QY	1981	GGGGCCTACTGTGAAACCAATATTAATGACTGACGACAGAACCCCTGCCCAAAATGGGGC	2040
Db	2244	GGGGCCTACTGTGAAACCAATATTAATGACTGACGACAGAACCCCTGCCCAAAATGGGGC	2303
QY	2041	ACGTGTGCGACCTGTGCAATGACTTCTACTGTGACTGTGTAATAATGGGTGGAAGGAAG	2100
Db	2304	ACGTGTGCGACCTGTGCAATGACTTCTACTGTGACTGTGTAATAATGGGTGGAAGGAAG	2363
QY	2101	ACCTGCCACTCACGTGACAGTCACTGTGATGAGGCCACGTGCAACACCGTGGCACCCTGC	2160
Db	2364	ACCTGCCACTCACGTGACAGTCACTGTGATGAGGCCACGTGCAACACCGTGGCACCCTGC	2423
QY	2161	TATGATGAGGGGATGCTTTTAAGTGCATGTGCTCTGGCGCTGGGAAGGAAACAACTGT	2220
Db	2424	TATGATGAGGGGATGCTTTTAAGTGCATGTGCTCTGGCGCTGGGAAGGAAACAACTGT	2483
QY	2221	AACATAGCCGAAACAGTACGTGCTGCCCAACCCCTGCCATTAATGGGGGACATGTGTG	2280
Db	2484	AACATAGCCGAAACAGTACGTGCTGCCCAACCCCTGCCATTAATGGGGGACATGTGTG	2543
QY	2281	GTCAACGGCGAGTCTTTTACGTGGTCTGCAAGGAAGGCTGGGAGGGGCCCATCTGTGCT	2340
Db	2544	GTCAACGGCGAGTCTTTTACGTGGTCTGCAAGGAAGGCTGGGAGGGGCCCATCTGTGCT	2603
QY	2341	CAGAATCAATGACTCGAGCCCTCATCCCTGTTTACAACAGCGGACCTGTGTGGATGGA	2400
Db	2604	CAGAATCAATGACTCGAGCCCTCATCCCTGTTTACAACAGCGGACCTGTGTGGATGGA	2663
QY	2401	GACAACTGGTACCGGTGGGATGTGCCCGGGTTTCTGGCGCCGACTCGAATAAAC	2460
Db	2664	GACAACTGGTACCGGTGGGATGTGCCCGGGTTTCTGGCGCCGACTCGAATAAAC	2723
QY	2461	ATCAATGAATGCCAGTCTTACCTTGTGCTTGGAGCGACCTGTGTGGATGAGATCAAT	2520
Db	2724	ATCAATGAATGCCAGTCTTACCTTGTGCTTGGAGCGACCTGTGTGGATGAGATCAAT	2783
QY	2521	GGCTACCGGTGTGTGCTCCCTCCAGGGCACAGTGGTGCCCAAGTGCCAGGAAGTTTCAGGG	2580
Db	2784	GGCTACCGGTGTGTGCTCCCTCCAGGGCACAGTGGTGCCCAAGTGCCAGGAAGTTTCAGGG	2843
QY	2581	AGACCTTGCATCACCATGGGGAGTGTGATACCAGATGGGGCCAAATGGGATGATGACTGT	2640
Db	2844	AGACCTTGCATCACCATGGGGAGTGTGATACCAGATGGGGCCAAATGGGATGATGACTGT	2903
QY	2641	AATACCTGCCAGTCCCTGAATGGACGATGCCCTGCTCAAAGGTCTGGTGGGCCCTCGA	2700
Db	2904	AATACCTGCCAGTCCCTGAATGGACGATGCCCTGCTCAAAGGTCTGGTGGGCCCTCGA	2963
QY	2701	CCTTGCTGTCTCCAAAGGGCACAGCGAGTGGCCCCAGCGGGCAGAGCTGATCCCCATC	2760
Db	2964	CCTTGCTGTCTCCAAAGGGCACAGCGAGTGGCCCCAGCGGGCAGAGCTGATCCCCATC	3023
QY	2761	CTGGAACGACGTGCTTGTCCACCTGCACTGTGTGGCGAGTGTGGTCTTCCAGT	2820
Db	3024	CTGGAACGACGTGCTTGTCCACCTGCACTGTGTGGCGAGTGTGGTCTTCCAGT	3083
QY	2821	CTCCAGCGGTGAAGACAAAGTGCACCTCTGACTCTTATTAACAGGATAACTGTGCGAAC	2880
Db	3084	CTCCAGCGGTGAAGACAAAGTGCACCTCTGACTCTTATTAACAGGATAACTGTGCGAAC	3143
QY	2881	ATCAATTTACTTTTAAACAGGAGATGATGTCAACAGGTCTTACTAGGAGCAATTTGC	2940
Db	3144	ATCAATTTACTTTTAAACAGGAGATGATGTCAACAGGTCTTACTAGGAGCAATTTGC	3203

QY	2941	AGTGAATTGAGGAATTTGAATATTTTGAAGAATGTTTCOGCTGAATATTCAATCTACATC	3000
Db	3204	AGTGAATTGAGGAATTTGAATATTTTGAAGAATGTTTCOGCTGAATATTCAATCTACATC	3263
QY	3001	GCTTGGAGCCTTCCCTTTCAGCGAACAAATGAATAATCATGTGGCCATTTCTGCTGAAGAT	3060
Db	3264	GCTTGGAGCCTTCCCTTTCAGCGAACAAATGAATAATCATGTGGCCATTTCTGCTGAAGAT	3323
QY	3061	ATACGGGATGATGGGAACCCGATCAAGGAATCACTGACAAATAATCGATCTTGTGTAGT	3120
Db	3324	ATACGGGATGATGGGAACCCGATCAAGGAATCACTGACAAATAATCGATCTTGTGTAGT	3383
QY	3121	AAACGTGATGGAACAGCTCGCTGATTGCTGCGTTGCGAAGTAAGAGTTCAGAGCGG	3180
Db	3384	AAACGTGATGGAACAGCTCGCTGATTGCTGCGTTGCGAAGTAAGAGTTCAGAGCGG	3443
QY	3181	CCTCTGAAGAACAAGACAGATTTCCTTGTTCCTCTGCTGAGCTCTGTCTTAACTGTGGCT	3240
Db	3444	CCTCTGAAGAACAAGACAGATTTCCTTGTTCCTCTGCTGAGCTCTGTCTTAACTGTGGCT	3503
QY	3241	TGATCTGTGCTTGTGACGGCTTCTACTGTGCTTGGGAGCGGGGAGAGCCGGG	3300
Db	3504	TGATCTGTGCTTGTGACGGCTTCTACTGTGCTTGGGAGCGGGGAGAGCCGGG	3563
QY	3301	AGCCACACACACTCAGCTCTGAGGACAAACACCAACAAACAACTGCGGGAGAGCTGAAC	3360
Db	3564	AGCCACACACACTCAGCTCTGAGGACAAACACCAACAAACAACTGCGGGAGAGCTGAAC	3623
QY	3361	CAGATCAAAAAACCCCATTTGAGAAAACATGGGGCCAAACACCGTCCCATCAAAGATTACGAG	3420
Db	3624	CAGATCAAAAAACCCCATTTGAGAAAACATGGGGCCAAACACCGTCCCATCAAAGATTACGAG	3683
QY	3421	AACAAGAACTCCAAATGCTTAATAATAGGACACACAACTTCTCAAGTGTAGAGAGGACGAC	3480
Db	3684	AACAAGAACTCCAAATGCTTAATAATAGGACACACAACTTCTCAAGTGTAGAGAGGACGAC	3743
QY	3481	ATGGAACAAACACCCAGAGAAAGCCCGTTTGGCAAGCAGCGCGGCTATAGCTGGTAGAC	3540
Db	3744	ATGGAACAAACACCCAGAGAAAGCCCGTTTGGCAAGCAGCGCGGCTATAGCTGGTAGAC	3803
QY	3541	AGAGAAGAGAAGCCCCCAACGACCGCCGACAAAAACCCAAACTGGAACAAACAAACAG	3600
Db	3804	AGAGAAGAGAAGCCCCCAACGACCGCCGACAAAAACCCAAACTGGAACAAACAAACAG	3863
QY	3601	GACAAACAGAGACTTGGAAAGTCCCGAGAGCTTAAACCGAATGGGTACATCGTATAG	3657
Db	3864	GACAAACAGAGACTTGGAAAGTCCCGAGAGCTTAAACCGAATGGGTACATCGTATAG	3920

RESULT 13

CQ727152 LOCUS CQ727152 DEFINITION Sequence 13086 from Patent WO02068579. CQ727152 ACCESSION CQ727152.1 GI:42292342 CQ727152.1 VERSION CQ727152.1

KEYWORDS Homo sapiens (human) SOURCE Homo sapiens ORGANISM Homo sapiens

REFERENCE 1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W. Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof

JOURNAL PE Corporation (NY) (US) Patent: WO 02068579-A 13086 06-SEP-2002;

FEATURES source location/Qualifiers 1. 5897 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"

ORIGIN

Query Match 99.9%; Score 3651; DB 6; Length 5897;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3651; Conservative 3; Mismatches 3; Indels 0

Qy	1	ATGCGTTCCCCACCGACRCGCGCGCGCGTTCGGGGCGCCCTTAAGACTTCCTGCTCGCCCTG	60
Db	413	ATGCGTTCCCCACCGACCGGCGCGCGTTCGGGGCGCCCTTAAGACTTCCTGCTCGCCCTG	472
Qy	61	CTCTGTGCCCTCGCAGACCAAGTGTGTGGGGCCTCGGGTCAAGTTCAGTTTCGAGTTGGAGATCCTG	120
Db	473	CTCTGTGCCCTCGCAGCCAAAGTGTGTGGGGCCTCGGGTCAAGTTCGAGTTTCGAGTTGGAGATCCTG	532
Qy	121	TCCATGCAGAACGTGAACGGGGAGCTGCAGAACCGGAACTGCTCGGGCGCGCCCGCAAC	180
Db	533	TCCATGCAGAACGTGAACGGGGAGCTGCAGAACCGGAACTGCTCGGGCGCGCCCGCAAC	592
Qy	181	CCGGGAGACCGCAAGTGCACCCCGCAGCAGAGTGTGACACATACTTCAAAGTGTGCTCAAG	240
Db	593	CCGGGAGACCGCAAGTGCACCCCGCAGCAGAGTGTGACACATACTTCAAAGTGTGCTCAAG	652
Qy	241	GAGTATCAGTCCCOCGTTCACGGCCGGGGCCCTGCAGCTTCGGCTCAGGGTCCACGCCT	300
Db	653	GAGTATCAGTCCCOCGTTCACGGCCGGGGCCCTGCAGCTTCGGCTCAGGGTCCACGCCT	712
Qy	301	GTCAATCGGGGGCAACAACCTTCAACCTCAAGGCCAGCCCGGCAACGACCGCAACCGCATC	360
Db	713	GTCAATCGGGGGCAACAACCTTCAACCTCAAGGCCAGCCCGGCAACGACCGCAACCGCATC	772
Qy	361	GTGCTGCCTTTTCAGTTTCGGCTGGCCAGAGTCTTATACGTTGCTTGTGAGGGCGTGGAT	420
Db	773	GTGCTGCCTTTTCAGTTTCGGCTGGCCAGAGTCTTATACGTTGCTTGTGAGGGCGTGGAT	832
Qy	421	TCCAGTAATGACACCGGTTCAACCTGACAGATTAATTGAAAAGGCTTCTCACTCGGGCATG	480
Db	833	TCCAGTAATGACACCGGTTCAACCTGACAGATTAATTGAAAAGGCTTCTCACTCGGGCATG	892
Qy	481	ATCAACCCGAGCGCGAGTGGCAGACGCTGAGCAGAACACGGGCGTTGCCCACTTTTGAG	540
Db	893	ATCAACCCGAGCGCGAGTGGCAGACGCTGAGCAGAACACGGGCGTTGCCCACTTTTGAG	952
Qy	541	TATCAGATCCGCGTGACCTGTGATGACTACTACTATGCTTTGGCTGYAATAGTTCTCG	600
Db	953	TATCAGATCCGCGTGACCTGTGATGACTACTACTATGCTTTGGCTGCAATAGTTCTCG	1012
Qy	601	CGCCCCAGAGATGACTTTCTTTTGGACACTATGCTGTCACAGAGTGCACAAACCTTGC	660
Db	1013	CGCCCCAGAGATGACTTTCTTTTGGACACTATGCTGTCACAGAGTGCACAAACCTTGC	1072
Qy	661	ATGGAAGGCTGGATGGGCCCGCGAATGTAAACAGAGCTAATTTGCCGACAAGGCTGAGTCT	720
Db	1073	ATGGAAGGCTGGATGGGCCCGCGAATGTAAACAGAGCTAATTTGCCGACAAGGCTGAGTCT	1132
Qy	721	AAGCATGGGTCTTTGCAAACTCCACAGTGACTGACAGGTGCCAGTAYGGCTGCGAAGGCTG	780
Db	1133	AAGCATGGGTCTTTGCAAACTCCACAGTGACTGACAGGTGCCAGTAYGGCTGCGAAGGCTG	1192
Qy	781	TACTGTGATAGTGCATCCACACCCGGGATGCGTCCACGGCATCTGTAAATGAGCCCTGG	840
Db	1193	TACTGTGATAGTGCATCCACACCCGGGATGCGTCCACGGCATCTGTAAATGAGCCCTGG	1252
Qy	841	CAGTGCCTCTGTGAGACCAACTGGGGCGGCACGCTCTGTGACAAAGATCTCAATTACTGT	900
Db	1253	CAGTGCCTCTGTGAGACCAACTGGGGCGGCACGCTCTGTGACAAAGATCTCAATTACTGT	1312
Qy	901	GGGACTCATCAGCCGTGCTCAACGGGGGAACTTGTAGCAACACAGGCCCTGACAAATAT	960
Db	1313	GGGACTCATCAGCCGTGCTCAACGGGGGAACTTGTAGCAACACAGGCCCTGACAAATAT	1372
Qy	961	CAGTGTCTTCGCCCTGAGGGGTAATTCAGGACCCAACTGTGAATTCGTAGACACGCTGC	1020
Db	1373	CAGTGTCTTCGCCCTGAGGGGTAATTCAGGACCCAACTGTGAATTCGTAGACACGCTGC	1432

Qy	1021	CTCTCTGATCCCTGTGTCAAAACAGAGGCAGCTGTAAAGAGACCTCCCTCGGGCTTTGAGTGT	1080
Db	1433	CTCTCTGATCCCTGTGTCAAAACAGAGGCAGCTGTAAAGAGAGACCTCCCTCGGGCTTTGAGTGT	1492
Qy	1081	GAGTGTTTCCCAAGCTCGACCGGGCCACATGCTCTACAAACATTTGATGACTGTTCTTCCT	1140
Db	1493	GAGTGTTTCCCAAGCTCGACCGGGCCACATGCTCTACAAACATTTGATGACTGTTCTTCCT	1552
Qy	1141	AATAACTGTTTCCCACGGGGCACTGCGCAGGACCTGGTTAAACGGATTTAAAGTGTGTGTGC	1200
Db	1553	AATAACTGTTTCCCACGGGGCACTGCGCAGGACCTGGTTAAACGGATTTAAAGTGTGTGTGC	1612
Qy	1201	CCCCACAGTGGACTGGGAAAACGTCGCCAGTTAGATGCAAAATGAATGTGAGGCCCAACCT	1260
Db	1613	CCCCACAGTGGACTGGGAAAACGTCGCCAGTTAGATGCAAAATGAATGTGAGGCCCAACCT	1672
Qy	1261	TGTGTAAACGCCAAATCCTGTAAAGAACTCTCATTTGCCAGCTACTACTCGCGACTGTCTTCCC	1320
Db	1673	TGTGTAAACGCCAAATCCTGTAAAGAACTCTCATTTGCCAGCTACTACTCGCGACTGTCTTCCC	1732
Qy	1321	GGCTGGATGGGTGAGAAATTTGTGACATAAAATATTAATGACTGCTTGGCCACTGTGCAGAAAT	1380
Db	1733	GGCTGGATGGGTGAGAAATTTGTGACATAAAATATTAATGACTGCTTGGCCACTGTGCAGAAAT	1792
Qy	1381	GAGCCCTCCTGTGGGATTTGGTTAATGGTTATCGCTGTATCTGTCCACCTGCGCTATGCA	1440
Db	1793	GAGCCCTCCTGTGGGATTTGGTTAATGGTTATCGCTGTATCTGTCCACCTGCGCTATGCA	1852
Qy	1441	GGCGATCACTGTGAGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGGT	1500
Db	1853	GGCGATCACTGTGAGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGGT	1912
Qy	1501	CACGTGCAGAAATGAAATCAAACAGATTCCAGTGTCTGTGTCCCACTGGTTTCTCTGGAAC	1560
Db	1913	CACGTGCAGAAATGAAATCAAACAGATTCCAGTGTCTGTGTCCCACTGGTTTCTCTGGAAC	1972
Qy	1561	CTCTGCAGCTGACATCGAATTTATGTGAGCTTAATCCCTGCCAGAACGGTGGCCAGTGC	1620
Db	1973	CTCTGCAGCTGACATCGAATTTATGTGAGCTTAATCCCTGCCAGAACGGTGGCCAGTGC	2032
Qy	1621	TACAAACCGTGCCAGTGACTATTTCTGCAAGTGCCTGAGGAGCTATGAGGCAAGAACTGC	1680
Db	2033	TACAAACCGTGCCAGTGACTATTTCTGCAAGTGCCTGAGGAGCTATGAGGCAAGAACTGC	2092
Qy	1681	TCACACTGTAAAGACCACTGCGCGGACGACCCCTCTGTGAAGTGAATGACAGCTGCACAGTG	1740
Db	2093	TCACACTGTAAAGACCACTGCGCGGACGACCCCTCTGTGAAGTGAATGACAGCTGCACAGTG	2152
Qy	1741	GCCATGGCTTCCAAACGACACACTGGAAGGGTGGGTATATTTTCTTCCACAGCTGTGGT	1800
Db	2153	GCCATGGCTTCCAAACGACACACTGGAAGGGTGGGTATATTTTCTTCCACAGCTGTGGT	2212
Qy	1801	CCTCAGGGGAAGTGCAGAGTCACTGGGAGGCAAAATTCACCTGTGACTGTATAACAAAGC	1860
Db	2213	CCTCAGGGGAAGTGCAGAGTCACTGGGAGGCAAAATTCACCTGTGACTGTATAACAAAGC	2272
Qy	1861	TTCAACGGGAACATACCTGCGCATGAAAAATATTAATGACTGTGAGAGCAACCCCTTGTAGAAC	1920
Db	2273	TTCAACGGGAACATACCTGCGCATGAAAAATATTAATGACTGTGAGAGCAACCCCTTGTAGAAC	2332
Qy	1921	GGTGGCACTTGCATCGATGGTGTCAACTCTACAGTGCATCTGTAGTACGGCTGGAG	1980
Db	2333	GGTGGCACTTGCATCGATGGTGTCAACTCTACAGTGCATCTGTAGTACGGCTGGAG	2392
Qy	1981	GGGGCCTACTGTGAAACCAATATTAATGACTGACGACGAAACCCCTCTGCCACATGGGGC	2040
Db	2393	GGGGCCTACTGTGAAACCAATATTAATGACTGACGACGAAACCCCTCTGCCACATGGGGC	2452
Qy	2041	ACGTGTGGGACCTGGTCAATGACTTCTACTGTGACTGTGATAAAATGGGTGGAAAGAAAG	2100
Db	2453	ACGTGTGGGACCTGGTCAATGACTTCTACTGTGACTGTGATAAAATGGGTGGAAAGAAAG	2512

```
QY 2101 ACCTGCCACTCAGTGACAGTCACTGATGAGGCCACGTGCAACACGGTGGCACTGC 2160
D 2513 ACCTGCCACTCAGTGACAGTCACTGATGAGGCCACGTGCAACACGGTGGCACTGC 2572
QY 2161 TATGATGAGGGGATGCTTTTAAAGTGATGTCTCTGGCGGCTGGGAAGGAACAACCTGT 2220
D 2573 TATGATGAGGGGATGCTTTTAAAGTGATGTCTCTGGCGGCTGGGAAGGAACAACCTGT 2632
QY 2221 AACATAGCCGGAACAAGTAGCTGCTGCCCAACCCCTGCCAATATGGGGGACATGTGTG 2280
D 2633 AACATAGCCGGAACAAGTAGCTGCTGCCCAACCCCTGCCAATATGGGGGACATGTGTG 2692
QY 2281 GTCAACGGCGAGTCTTTTAAAGTGATGTCTCTGGCGGCTGGGAAGGAACAACCTGT 2340
D 2693 GTCAACGGCGAGTCTTTTAAAGTGATGTCTCTGGCGGCTGGGAAGGAACAACCTGT 2752
QY 2341 CAGAATACCAATGACTGCGACCTCATCCCTGTTTACAAACAGCGGACCTGTGTGGATGGA 2400
D 2753 CAGAATACCAATGACTGCGACCTCATCCCTGTTTACAAACAGCGGACCTGTGTGGATGGA 2812
QY 2401 GACAACTGGTACCGGTGGGAATGTGCGCGGCTTTTCTGGGCGCCGACCTGCGAGATTAAC 2460
D 2813 GACAACTGGTACCGGTGGGAATGTGCGCGGCTTTTCTGGGCGCCGACCTGCGAGATTAAC 2872
QY 2461 ATCAATGAATGCCAGTCTTCACTTGTGCTTTGGAGGACCTGTGTGGATGAGATCAAT 2520
D 2932 ATCAATGAATGCCAGTCTTCACTTGTGCTTTGGAGGACCTGTGTGGATGAGATCAAT 2932
QY 2521 GGCTACCGGTGTGCTGCGCTTCCAGGCGCACAGTGGTGCCTCAAGTCCAGGGAAGTTTCAGGG 2580
D 2933 GGCTACCGGTGTGCTGCGCTTCCAGGCGCACAGTGGTGCCTCAAGTGGCCAGGGAAGTTTCAGGG 2992
QY 2581 AGACCTTGCAATCACTGCGGAGTGTGATACCAAGTGGGGGCAAAATGGGATGATGATGT 2640
D 3052 AGACCTTGCAATCACTGCGGAGTGTGATACCAAGTGGGGGCAAAATGGGATGATGATGT 3052
QY 2641 AATACCTGCCAGTGCCTGAAATGGAGCGGATGCCCTGCTCAAGGTCTGCTGGGCGCCCTGA 2700
D 3112 AATACCTGCCAGTGCCTGAAATGGAGCGGATGCCCTGCTCAAGGTCTGCTGGGCGCCCTGA 3112
QY 2701 CTTGCTGCTTCCAAAGGGCACAGGAGTGTGCTGCGGCGGCGAGAGTGCATCCCCATC 2760
D 3113 CTTGCTGCTTCCAAAGGGCACAGGAGTGTGCTGCGGCGGCGAGAGTGCATCCCCATC 3172
QY 2761 CTGAGCACAGTGTGCTTCCACCCCTGCACTGGTGTGGCGGAGTGTGCTGCTTCCAGT 2820
D 3173 CTGAGCACAGTGTGCTTCCACCCCTGCACTGGTGTGGCGGAGTGTGCTGCTTCCAGT 3232
QY 2821 CTCCAGCGGTGAAGACAAAGTGCACCTCTGACTCTTATACCAAGGATACTGTGCGAAC 2880
D 3233 CTCCAGCGGTGAAGACAAAGTGCACCTCTGACTCTTATACCAAGGATACTGTGCGAAC 3292
QY 2881 ATCACTTTACCTTTAAACAGGAGATGATGTCACCAAGTCTTACTACGGACACATTTGC 2940
D 3293 ATCACTTTACCTTTAAACAGGAGATGATGTCACCAAGTCTTACTACGGACACATTTGC 3352
QY 2941 AGTGAATGAGGAATTTGAAATTTTGAAGAAATGTTTCGCTGAAATTTCAATCTACATC 3000
D 3412 AGTGAATGAGGAATTTGAAATTTTGAAGAAATGTTTCGCTGAAATTTCAATCTACATC 3412
QY 3001 GCTTGGAGCTTCCCTTTCAGCGAACCAATGAATACATGATGGCCATTTCTGCTGAAGAT 3060
D 3413 GCTTGGAGCTTCCCTTTCAGCGAACCAATGAATACATGATGGCCATTTCTGCTGAAGAT 3472
QY 3061 ATACGGGATGATGGGAACCCGATCAAGGAAATCACTGACAAAATAATCGATCTTTGTAGT 3120
D 3473 ATACGGGATGATGGGAACCCGATCAAGGAAATCACTGACAAAATAATCGATCTTTGTAGT 3532
QY 3121 AAACTGATGAAACAGCTGCTGATGCTGCGGTGAGAAATGAGTTTCAGAGCGG 3180
D 3533 AAACTGATGAAACAGCTGCTGATGCTGCGGTGAGAAATGAGTTTCAGAGCGG 3592
QY 3181 CCTCTGAAGAACAGAAACAGATTTCTTCTTCTGCTGAGCTCTGCTTAACTGTGCT 3240
```

```
Db 3593 CCTCTGAAGAACAGAACAGATTTCTTCTTCCCTTCTGCTGAGCTCTGTCTTAACTGTGGCT 3652
QY 3241 TGGATCTGTTGCTTGTGTGAGCGCTTCTACTGTGTGCTGCGGAAGCGGCGAAGCGGGC 3300
D 3653 TGGATCTGTTGCTTGTGTGAGCGCTTCTACTGTGTGCTGCGGAAGCGGCGAAGCGGGC 3712
QY 3301 AGCCACACACACTCAGCTCTGAGGAACAACACCAACAAACAAAGTGGGAGAGCTGAAC 3360
D 3713 AGCCACACACACTCAGCTCTGAGGAACAACACCAACAAACAAAGTGGGAGAGCTGAAC 3772
QY 3361 CAGATCAAAAAACCCCAATTTGAGAAACATGGGGCCAAACACCGTCCCCATCAAGGATTACGAG 3420
D 3773 CAGATCAAAAAACCCCAATTTGAGAAACATGGGGCCAAACACCGTCCCCATCAAGGATTACGAG 3832
QY 3421 AACAGAACTCCAAAATGTCTAAAATAAGGACACAAATTTCTGAAGTAGAAGAGGACGAC 3480
D 3833 AACAGAACTCCAAAATGTCTAAAATAAGGACACAAATTTCTGAAGTAGAAGAGGACGAC 3892
QY 3481 ATGGACAAACACACAGCAGAAAGCCCGTTTGGCAAGCAGCGCGCTATACGCTGGTAGAC 3540
D 3893 ATGGACAAACACACAGCAGAAAGCCCGTTTGGCAAGCAGCGCGCTATACGCTGGTAGAC 3592
QY 3541 AGAGAAGAGAAAGCCCGCAACCGCACCGGCAAAAAACCCAAACTGGACAAACAAACAG 3600
D 3953 AGAGAAGAGAAAGCCCGCAACCGCACCGGCAAAAAACCCAAACTGGACAAACAAACAG 4012
QY 3601 GACAAACAGAGACTTGGAAAGTCCCGAGAGCTTAAACCGAATGAGTACATCGTATAG 3657
D 4013 GACAAACAGAGACTTGGAAAGTCCCGAGAGCTTAAACCGAATGAGTACATCGTATAG 4069

RESULT 14
AR482475
LOCUS AR482475 4855 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 34 from patent US 6703204.
ACCESSION AR482475
VERSION AR482475.1 GI:47244908
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4855)
AUTHORS Muter, G. L. and Baek, J. P. A.
TITLE Prognostic classification of breast cancer through determination of
nucleic acid sequence expression
JOURNAL Patent: US 6703204-A 34 09-MAR-2004;
The Brigham & Women's Hospital, Inc.; Boston, MA
FEATURES
source Location/Qualifiers
1..4855
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 99.8%; Score 3649.4; DB 6; Length 4855;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3650; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGCGTTCCCAACGACACACCGCGCGCTGCGGGCGCCCTTAAGCCCTCTGCTGCGCCCTG 60
D 371 ATGCGTTCCCAACGACACACCGCGCGCTGCGGGCGCCCTTAAGCCCTCTGCTGCGCCCTG 430
QY 61 CTCTGTGCTTCCGAGCAAGAGTGTGTGGGGCCCTGCGGCTCAGTTCGAGTTGGAGATCTG 120
D 431 CTCTGTGCTTCCGAGCAAGAGTGTGTGGGGCCCTGCGGCTCAGTTCGAGTTGGAGATCTG 490
QY 121 TCCATGACAGAACTGAAACGGGAGCTGCAGAAACGGGAACTGCTGCGGCGCGCCGGAAAC 180
D 491 TCCATGACAGAACTGAAACGGGAGCTGCAGAAACGGGAACTGCTGCGGCGCGCCGGAAAC 550
QY 181 CCGGAGACCGCAAGTGCACCCCGCGACAGTGTGACACATCTTCAAAGTGTGCTCTCAG 240
D 551 CCGGAGACCGCAAGTGCACCCCGCGACAGTGTGACACATCTTCAAAGTGTGCTCTCAG 610
```

Qy	241	GAGTATCAGTCCCGCGTCA	CGGCCGGGGGGCCCTG	CAGCTTCGGCTCAGGGTCCACGCCT	300
Db	611	GAGTATCAGTCCCGCGT	CAGCGCCGGGGGGCCCTG	CAGCTTCGGCTCAGGGTCCACGCCT	670
Qy	301	GTCAATCGGGGCAACACCTT	CAACCTCAAGGCGAGCGCGGCAACGACCGCATC	360	
Db	671	GTCAATCGGGGCAACACCTT	CAACCTCAAGGCGAGCGCGGCAACGACCGCATC	730	
Qy	361	GTGCTGCCTTTCA	GTTTGCCTGCGCGAGGFTCTATACGTTGCTTGTGGAGGCGTGGGAT	420	
Db	731	GTGCTGCCTTTCA	GTTTGCCTGCGCGAGGFTCTATACGTTGCTTGTGGAGGCGTGGGAT	790	
Qy	421	TCCAGTAATGACACCGTT	CAACCTGACAGTATTATTGAAAAGGCTTCTCACTCGGGCATG	480	
Db	791	TCCAGTAATGACACCGTT	CAACCTGACAGTATTATTGAAAAGGCTTCTCACTCGGGCATG	850	
Qy	481	ATCAACCCCGAGCCGCGAGT	GGCAGACGCTGAAGCAGAACACCGGCGCTTGCCCACTTTGAG	540	
Db	851	ATCAACCCCGAGCCGCGAGT	GGCAGACGCTGAAGCAGAACACCGGCGCTTGCCCACTTTGAG	910	
Qy	541	TATCAGATCCGCGTGAC	CTGTGATGACTACTATATGCGCTTTGCGCTGYAATAAGTTCTGC	600	
Db	911	TATCAGATCCGCGTGAC	CTGTGATGACTACTATATGCGCTTTGCGCTGYAATAAGTTCTGC	970	
Qy	601	CGCCCCAGAGATGACTT	CTTTGGACACTATGCGCTGTACCCAGTAATGCAACAAACTTGC	660	
Db	971	CGCCCCAGAGATGACTT	CTTTGGACACTATGCGCTGTACCCAGTAATGCAACAAACTTGC	1030	
Qy	661	ATGGAAGGCTGGATGGG	CCCCCGAATTTGCCGACAAAGGCTGCAGTCCCT	720	
Db	1031	ATGGAAGGCTGGATGGG	CCCCCGAATTTGCCGACAAAGGCTGCAGTCCCT	1090	
Qy	721	AAGCATGGGCTTTC	CAAACTCCAGGTGACTGCAAGGTGCCAGTACGGCTGGCAAGGCGCTG	780	
Db	1091	AAGCATGGGCTTTC	CAAACTCCAGGTGACTGCAAGGTGCCAGTACGGCTGGCAAGGCGCTG	1150	
Qy	781	TACTGTGATAAGTGCAT	CCACACCCGGGATGGCTCCACGGCATCTGTATGATGAGCCCTGG	840	
Db	1151	TACTGTGATAAGTGCAT	CCACACCCGGGATGGCTCCACGGCATCTGTATGATGAGCCCTGG	1210	
Qy	841	CAGTGCCTCTGTGAGACA	CAACTGGGGCGGCAGCTCTGTGACAAAGATCTCAATTA	900	
Db	1211	CAGTGCCTCTGTGAGACA	CAACTGGGGCGGCAGCTCTGTGACAAAGATCTCAATTA	1270	
Qy	901	GGGACTCATCAGCCGTGCT	CAACGGGGGAACTTGATGACAAACAGGCCCTGACAAATAT	960	
Db	1271	GGGACTCATCAGCCGTGCT	CAACGGGGGAACTTGATGACAAACAGGCCCTGACAAATAT	1330	
Qy	961	CAGTGTTCCTGCCCTGAGGGGT	ATTACAGACCCAACTGTCAAAATTCGTGACGACGCTGC	1020	
Db	1331	CAGTGTTCCTGCCCTGAGGGGT	ATTACAGACCCAACTGTGAAAATTCGTGACGACGCTGC	1390	
Qy	1021	CTCTCTGATCCCTGTGCA	CAACAGAGGCAGCTGTAAAGGAGACCTCCCTGGGCTTTGAGTGT	1080	
Db	1391	CTCTCTGATCCCTGTGCA	CAACAGAGGCAGCTGTAAAGGAGACCTCCCTGGGCTTTGAGTGT	1450	
Qy	1081	GAGTGTTCCTCCAGGCTGGA	CGGGCCCCACATGCTCTACAAACATTGATGACTGTTCTCCT	1140	
Db	1451	GAGTGTTCCTCCAGGCTGGA	CGGGCCCCACATGCTCTACAAACATTGATGACTGTTCTCCT	1510	
Qy	1141	AATAACTGTTCCCAACGGGGG	CACCTGCGAGGACTGGTTAAACGGAATTAAGTGTGTGTGC	1200	
Db	1511	AATAACTGTTCCCAACGGGGG	CACCTGCGAGGACTGGTTAAACGGAATTAAGTGTGTGTGC	1570	
Qy	1201	CCCCCAGGTGACTGGGAAA	ACGTGCCAGTTTGAATGCAAAATGAATGTGAGGCCAAACCT	1260	
Db	1571	CCCCCAGGTGACTGGGAAA	ACGTGCCAGTTTGAATGCAAAATGAATGTGAGGCCAAACCT	1630	
Qy	1261	TGTGTAAACGCCAAATCCT	GTAAAGATCTCAATTGCCAGCTACTACTGCGACTGTCTTCCC	1320	
Db	1631	TGTGTAAACGCCAAATCCT	GTAAAGATCTCAATTGCCAGCTACTACTGCGACTGTCTTCCC	1690	

Qy	1321	GGCTGGATGGGT	CAGAAATTGTGTGACATAAATATTAATGACTGCCTTGGCCACGTGT	CAGAAAT	1381
Db	1691	GGCTGGATGGGT	CAGAAATTGTGTGACATAAATATTAATGACTGCCTTGGCCACGTGT	CAGAAAT	1750
Qy	1381	GAGCCCTCCTGT	CGGGATTTGGTTAATGGTTATCGCTGTATCTGTGCCACCTGGCTATGCA	1440	
Db	1751	GAGCCCTCCTGT	CGGGATTTGGTTAATGGTTATCGCTGTATCTGTCCACTGGCTATGCA	1810	
Qy	1441	GGCGATCACTGT	GAGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAAATGGGGGT	1500	
Db	1811	GGCGATCACTGT	GAGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAAATGGGGGT	1870	
Qy	1501	CACGTGCAGATGAAAT	CAACAGATTCAGAGTGTCTGTGCCAGACGGTGTCTCTGGAAAC	1560	
Db	1871	CACGTGCAGATGAAAT	CAACAGATTCAGAGTGTCTGTGCCAGACGGTGTCTCTGGAAAC	1930	
Qy	1561	CTCTGT	CAGCTGGACATCGATTAATGTGAGCCTAAATCCCTGCCAGAACGGTGC	1620	
Db	1931	CTCTGT	CAGCTGGACATCGATTAATGTGAGCCTAAATCCCTGCCAGAACGGTGC	1990	
Qy	1621	TACAACCGTGCCAGT	GACTATTTCTGCAAGTGCCCGAGGACTATGAGGGCAAGAACTGC	1680	
Db	1991	TACAACCGTGCCAGT	GACTATTTCTGCAAGTGCCCGAGGACTATGAGGGCAAGAACTGC	2050	
Qy	1681	TCACACCTGAAAGACCA	CTGCCGACGACCCCTGTGAAGTGATGACAGCTGCACAGTG	1740	
Db	2051	TCACACCTGAAAGACCA	CTGCCGACGACCCCTGTGAAGTGATGACAGCTGCACAGTG	2110	
Qy	1741	GCCATGGCTTCCAA	CGACACACCTGGAAGGGGTGCGGTATATTTCTTCCACACGTCTGTGGT	1800	
Db	2111	GCCATGGCTTCCAA	CGACACACCTGGAAGGGGTGCGGTATATTTCTTCCACACGTCTGTGGT	2170	
Qy	1801	CCTCACGGGAAGTGC	CAAGAGTCAAGTGGGAGGCAAAATTCACCTGTGACTGTAAACAAGGC	1860	
Db	2171	CCTCACGGGAAGTGC	CAAGAGTCAAGTGGGAGGCAAAATTCACCTGTGACTGTAAACAAGGC	2230	
Qy	1861	TTCAACGGGAACATAC	TGCGCATGAAAATATTAATGACTGTGAGAGCAACCCCTGTAGAAAC	1920	
Db	2231	TTCAACGGGAACATAC	TGCGCATGAAAATATTAATGACTGTGAGAGCAACCCCTGTAGAAAC	2290	
Qy	1921	GGTGGCACTTTGCAT	CGATGTGTCAACTCCTACAAGTGCACTGTGAGAGCAACCCCTGTGGAG	1980	
Db	2291	GGTGGCACTTTGCAT	CGATGTGTCAACTCCTACAAGTGCACTGTGAGAGCAACCCCTGTGGAG	2350	
Qy	1981	GGGGCCTTACTGT	GAAACCAATATTAATGACTGCAGCGCAAGAACCCCTGCCACAATGGGGGC	2040	
Db	2351	GGGGCCTTACTGT	GAAACCAATATTAATGACTGCAGCGCAAGAACCCCTGCCACAATGGGGGC	2410	
Qy	2041	ACGTGTGCGCACTGT	CAATGACTTCTACTGTGACTGTAAAAATGGGTGGAAGGAAG	2100	
Db	2411	ACGTGTGCGCACTGT	CAATGACTTCTACTGTGACTGTAAAAATGGGTGGAAGGAAG	2470	
Qy	2101	ACCTGCCACTCACGT	GCACGTGATGAGGCCACGTCGCAACACCGTGGCCACTGC	2160	
Db	2471	ACCTGCCACTCACGT	GCACGTGATGAGGCCACGTCGCAACACCGTGGCCACTGC	2530	
Qy	2161	TATGATGAGGGGGAT	GCTTTTAAAGTGCATGTGCTCTGGGGCTGGGAAGGAACAACTGT	2220	
Db	2531	TATGATGAGGGGGAT	GCTTTTAAAGTGCATGTGCTCTGGGGCTGGGAAGGAACAACTGT	2590	
Qy	2221	AAATAGCCCGAACA	GACTGCTGCCCAACCCCTGCAATATGGGGGCAATGTGT	2280	
Db	2591	AAATAGCCCGAACA	GACTGCTGCCCAACCCCTGCAATATGGGGGCAATGTGT	2650	
Qy	2281	GTCAACGGCGAGT	CTTTTACGTGCGTCTGCAAGGAAGGCTGGGAGGGGCCCATCTGTGCT	2340	
Db	2651	GTCAACGGCGAGT	CTTTTACGTGCGTCTGCAAGGAAGGCTGGGAGGGGCCCATCTGTGCT	2710	
Qy	2341	CAGAATACCAATGACT	GCAGCCCTCATCCCTGTGTACAAACAGCGGCACCTGTGTGGATGGA	2400	
Db	2711	CAGAATACCAATGACT	GCAGCCCTCATCCCTGTGTACAAACAGCGGCACCTGTGTGGATGGA	2770	
Qy	2401	GACAACTGGTAC	CGGTGCGAATGTGCCCGGGTTTTTCTGGGGCCGACTGCGAATAAAC	2460	

Db 2771 GACAACTGGTACCGGTGCGAATGTGCCCGGGTTTGTCTGGGCCGAGCTGCAGATTAAC 2830
Qy 2461 ATCAATGAATGCCAGTCTTCACTTGTGCTTTGGAGCGACCTGTGTGGATGATCAAT 2520
Db 2831 ATCAATGAATGCCAGTCTTCACTTGTGCTTTGGAGCGACCTGTGTGGATGATCAAT 2890
Qy 2521 GGCTACCGGTGTGCTGCCCTCCAGGGCACAGTGGTCCCAAGTCCAGGAAGTTTCAGGG 2580
Db 2891 GGCTACCGGTGTGCTGCCCTCCAGGGCACAGTGGTCCCAAGTCCAGGAAGTTTCAGGG 2950
Qy 2581 AGACTTGTGATCACCATGGGAGTGTGATACAGATGGGCCCAAAATGGGATGATGACTGT 2640
Db 2951 AGACTTGTGATCACCATGGGAGTGTGATACAGATGGGCCCAAAATGGGATGATGACTGT 3010
Qy 2641 AATACCTGCCAGTCCCTGAATGAGACGGATGCCCTGTCTCAAGGTCTGGTGTGGCCCTCGA 2700
Db 3011 AATACCTGCCAGTCCCTGAATGAGACGGATGCCCTGTCTCAAGGTCTGGTGTGGCCCTCGA 3070
Qy 2701 CTTGCTGTCTCCAAAGGGCACAGGAGTGGCCCGAGCGGGCGAGCTGCATCCCATC 2760
Db 3071 CTTGCTGTCTCCAAAGGGCACAGGAGTGGCCCGAGCGGGCGAGCTGCATCCCATC 3130
Qy 2761 CTGACGACCAAGTCTTCTGCTCCACCTCTGCACTGTGTGGCGAGTGTCTGCTTCCAGT 2820
Db 3131 CTGACGACCAAGTCTTCTGCTCCACCTCTGCACTGTGTGGCGAGTGTCTGCTTCCAGT 3190
Qy 2821 CTCAGCGGTGGAAGACAAAGTGCACCTCTGACTCTATTACCAAGGATAAATCTGCGAAC 2880
Db 3191 CTCAGCGGTGGAAGACAAAGTGCACCTCTGACTCTATTACCAAGGATAAATCTGCGAAC 3250
Qy 2881 ATCAATTTACCTTTAAACAGGAGATGATGTCACCGGTCTTACTAGCGAGCATTTGC 2940
Db 3251 ATCAATTTACCTTTAAACAGGAGATGATGTCACCGGTCTTACTAGCGAGCATTTGC 3310
Qy 2941 AGTGAATGAGGAATTTGAATATTTGAAGAAATGTTTCCGCTGAATATTTCAATCTACATC 3000
Db 3311 AGTGAATGAGGAATTTGAATATTTGAAGAAATGTTTCCGCTGAATATTTCAATCTACATC 3370
Qy 3001 GCTTGCAGCGCTTCCCTTTAGCGAAACAAATGAATAATATGATGCGCATTTCTGCTGAAGAT 3060
Db 3371 GCTTGCAGCGCTTCCCTTTAGCGAAACAAATGAATAATATGATGCGCATTTCTGCTGAAGAT 3430
Qy 3061 ATAGCGATGATGGAAACCGCATCAAGGAATCACTGACAAATAATCGATCTTGTGTAGT 3120
Db 3431 ATA CGGATGATGGAAACCGCATCAAGGAATCACTGACAAATAATCGATCTTGTGTAGT 3490
Qy 3121 AAACGTGATGGAAACAGCTCGCTGATGCTGCCGTTGCAGAAATGAAGATTTCAAGCGCG 3180
Db 3491 AAACGTGATGGAAACAGCTCGCTGATGCTGCCGTTGCAGAAATGAAGATTTCAAGCGCG 3550
Qy 3181 CCTCTGAAGAACAGAACAGATTTCTTGTTCCTTGTGTGAGCTCTGTCTTAATCTGTGCT 3240
Db 3551 CCTCTGAAGAACAGAACAGATTTCTTGTTCCTTGTGTGAGCTCTGTCTTAATCTGTGCT 3610
Qy 3241 TGGATCTGTTGTTGGTGAACGCTTCTACTGTTGCTGCGGAAGCGCGGAAGCGCGGC 3300
Db 3611 TGGATCTGTTGTTGGTGAACGCTTCTACTGTTGCTGCGGAAGCGCGGAAGCGCGGC 3670
Qy 3301 AGCCACACACTCAGCTCTGAGGACAAACACCAACACAGTGGCGGAGCAGCTGAAC 3360
Db 3671 AGCCACACACTCAGCTCTGAGGACAAACACCAACACAGTGGCGGAGCAGCTGAAC 3730
Qy 3361 CAGATCAAAAAACCCCATTTAGAAACATGGGGCCAAACAGCGTCCCATCAAGATTAAGAG 3420
Db 3731 CAGATCAAAAAACCCCATTTAGAAACATGGGGCCAAACAGCGTCCCATCAAGATTAAGAG 3790
Qy 3421 AACAGAACTCCAAATCTTAATATAGGACACACAAATCTGAAGTAGAGAGACGAC 3480
Db 3791 AACAGAACTCCAAATCTTAATATAGGACACACAAATCTGAAGTAGAGAGACGAC 3850
Qy 3481 ATGGAACAAACCCAGCAGAAAGCCCGGTTTGGCAGCAGCGCGGTATATACCTGCTAGAC 3540

Db 3851 ATGGACAAACACCAGCAGAAAGCCCGGTTTGGCAAGCAGCGCGGTACACGCTGGTAGAC 3910
Qy 3541 AG 3600
Db 3911 AG 3970
Qy 3601 GACAAACAG 3657
Db 3971 GACAAACAG 4027

RESULT 15
AX375031
LOCUS AX375031 4855 bp DNA linear PAT 01-MAR-2002
DEFINITION Sequence 34 from Patent WO0210436.
ACCESSION AX375031
VERSION AX375031.1 GI:19169863
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

REFERENCE 1
AUTHORS Baak,J. and Mutter,G.L.
TITLE Prognostic classification of breast cancer
JOURNAL Patent: WO 0210436-A, 34, 07-FEB-2002;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US) ; Baak, Jan (US)

FEATURES
Location/Qualifiers
source 1..4855
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.8%; Score 3649.4; DB 6; Length 4855;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3650; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGCGTTCCCAACGAGACRCCGCGCGGTCGCGGGCGCCCTTAAAGCTTCTGCTCGCCCTG 60
Db 371 ATGCGTTCCCAACGAGACACGCGCGCGGTCGCGGGCGCCCTTAAAGCTTCTGCTCGCCCTG 430
Qy 61 CTCTGTGCTTCCGAGCAAGGTGTGTGGGCGCTCGGGCTCAGTTTCAAGTTGGAGATCCTG 120
Db 431 CTCTGTGCTTCCGAGCAAGGTGTGTGGGCGCTCGGGCTCAGTTTCAAGTTGGAGATCCTG 490
Qy 121 TCCATGACAGAACGTGAACGGGAGCTGCAGAACGGGAACTGCTGCGCGCGCGCCCGGAAC 180
Db 491 TCCATGACAGAACGTGAACGGGAGCTGCAGAACGGGAACTGCTGCGCGCGCGCCCGGAAC 550
Qy 181 CCGGAGACCGCAAGTGCACCCCGGAGAGTGTGACACATCTTCAAGTTGCGCTCAAG 240
Db 551 CCGGAGACCGCAAGTGCACCCCGGAGAGTGTGACACATCTTCAAGTTGCGCTCAAG 610
Qy 241 GAGTATCAGTCCCGCTCAGCGCGCGGCGCCCTGAGCTTCCGCTCAGGGTCCACGCT 300
Db 611 GAGTATCAGTCCCGCTCAGCGCGCGGCGCCCTGAGCTTCCGCTCAGGGTCCACGCT 670
Qy 301 GTCATCGGGGGCAACACCTTTCAACCTCAAGGCCAGCGCGCAACAGACCGCAACCGCATC 360
Db 671 GTCATCGGGGGCAACACCTTTCAACCTCAAGGCCAGCGCGCAACAGACCGCATC 730
Qy 361 GTCTGCTCTTTCAGTTTCGCTGGCCGAGTCTCTATAGTTGCTTGGAGGCGCTGGAT 420
Db 731 GTCTGCTCTTTCAGTTTCGCTGGCCGAGTCTCTATAGTTGCTTGGAGGCGCTGGAT 790
Qy 421 TCCAGTAATGACACCGGTTCAACCTGACAGTATTATTGAAAGGCTTCTCACTCGGGCATG 480
Db 791 TCCAGTAATGACACCGGTTCAACCTGACAGTATTATTGAAAGGCTTCTCACTCGGGCATG 850
Qy 481 ATCAACCCAGCGCGAGTGGCAGACCGCTGAAGCAGAACCGGGCGTTGCCCATCTTGAG 540

851	ATCAACCCAGCCGGCAGTGGCAGACCGTGAACAGAAACACGGGCGTTGCCCACTTTGAG	910	
Qy	541	TATCAGATCCGCGTGAACCTGTGATGACTACTACTATGGCTTTGGCTGYAATAAGTCTTCG	600
	911	TATCAGATCCGCGTGACCTGTGATGACTACTACTATGGCTTTGGCTTGTATAAGTCTTCG	970
	601	CGCCCCAGAGATGACTTCTTTTGGACACTATGCCCTGTGACCGAATGGCAACAAACTTGC	660
Db	971	CGCCCCAGAGATGACTTCTTTTGGACACTATGCCCTGTGACCGAATGGCAACAAACTTGC	1030
	661	ATGGAAGCTCGATGGGCCCGAATGTAAACAGAGCTATTTGCCGACAAGGCTGCAGTCCCT	720
	1031	ATGGAAGCTGATGGGCCCGAATGTAAACAGAGCTATTTGCCGACAAGGCTGCAGTCCCT	1090
Qy	721	AAGCATGGGCTTTCAAACTCCCAAGGTGACTGCAAGGTGCCAGTAYGGCTGGCAAGGCTTG	780
Db	1091	AAGCATGGGCTTTCAAACTCCCAAGGTGACTGCAAGGTGCCAGTAYGGCTGGCAAGGCTTG	1150
	781	TACTGTGATAAGTGCATCCCAACCCGGGATGCGTCCACGGCATCTCTATAAGAGCCCTGG	840
	1151	TACTGTGATAAGTGCATCCCAACCCGGGATGCGTCCACGGCATCTCTATAAGAGCCCTGG	1210
Qy	841	CAGTGCCCTCTGTGAGACCAACTGGGGGGCGCCAGCTCTGTGCAAGAATCTCAATTACTGT	900
Db	1211	CAGTGCCCTCTGTGAGACCAACTGGGGGGCGCCAGCTCTGTGCAAGAATCTCAATTACTGT	1270
	901	GGGACTCATCAGCCGCTGTCTCAACGGGGGAACTTTGTAGCAACACAGGCCCTGCAAAATAT	960
	1271	GGGACTCATCAGCCGCTGTCTCAACGGGGGAACTTTGTAGCAACACAGGCCCTGCAAAATAT	1330
Qy	961	CAGTGTTCTGCTCGCTGAGGGGTATTCAAGACCAACTGTGAAATTTGTGAGCAGCCCTGC	1020
Db	1331	CAGTGTTCTGCTCGCTGAGGGGTATTCAAGACCAACTGTGAAATTTGTGAGCAGCCCTGC	1390
	1021	CTCTCTCATCCCTGCTCAACACAGAGGCAGCTGTGAAGCAGACCTCCCTGGGCTTTGAGTGT	1080
	1391	CTCTCTCATCCCTGCTCAACACAGAGGCAGCTGTGAAGCAGACCTCCCTGGGCTTTGAGTGT	1450
Qy	1081	GAGTGTTCCCAGGCTGGACCGGCCCCACATGCTCTACAAACATTGATGACTGTTCTTCCT	1140
Db	1451	GAGTGTTCCCAGGCTGGACCGGCCCCACATGCTCTACAAACATTGATGACTGTTCTTCCT	1510
	1141	AATAACTGTTCCACCGGGGCACTGCGCAGGACCTGCTTTAACCGAATTAAGTGTGTGTGC	1200
	1511	AATAACTGTTCCACCGGGGCACTGCGCAGGACCTGCTTTAACCGAATTAAGTGTGTGTGC	1570
Qy	1201	CCCCACAGTGGACTGGGAAAACCTGCCATTTAGATGCAATGAATGAATGAGGCCCAACCT	1260
Db	1571	CCCCACAGTGGACTGGGAAAACCTGCCATTTAGATGCAATGAATGAATGAGGCCCAACCT	1630
	1261	TGTGTAAACGCCAAATCCTGTAAAGAACTCTCATTTGCCAGCTACTACTCGCACTGTCTTCCC	1320
	1631	TGTGTAAACGCCAAATCCTGTAAAGAACTCTCATTTGCCAGCTACTACTCGCACTGTCTTCCC	1690
Qy	1321	GGCTGGATGGGTGAGAAATTTGACATAAAATTAATGACTGCCCTTGGCCAGTGTGAGAAT	1380
Db	1691	GGCTGGATGGGTGAGAAATTTGACATAAAATTAATGACTGCCCTTGGCCAGTGTGAGAAT	1750
	1381	GACGCTCTCTGCGGATTTGGTTAAATGGTTATCGCTGTATCTGTCCACCTGGCATGCA	1440
	1751	GACGCTCTCTGCGGATTTGGTTAAATGGTTATCGCTGTATCTGTCCACCTGGCATGCA	1810
Qy	1441	GGCGATCACTGTGAGAGAGACATCGATGAATTTGGCCAGCAACCCCTCTTTGAAATGGGGGT	1500
Db	1811	GGCGATCACTGTGAGAGAGACATCGATGAATTTGGCCAGCAACCCCTCTTTGAAATGGGGGT	1870
	1501	CACTGTGAGAAATGAAATCAACAGATTCCAGTGTCTGTGTCCCACTGGTTCTCTGGAAC	1560
	1871	CACTGTGAGAAATGAAATCAACAGATTCCAGTGTCTGTGTCCCACTGGTTCTCTGGAAC	1930
Qy	1561	CTCTGTGAGTGCATCGAATTATTGTGAGCCTAAATCCCTGGCCAGAACGGTGCCCACTGC	1620
Db	1931	CTCTGTGAGTGCATCGAATTATTGTGAGCCTAAATCCCTGGCCAGAACGGTGCCCACTGC	1990

Qy 2701 CTTTGCCCTGCTCCAAAGGSCACAGCGAGTGCCCGAGCGGAGAGCTGATCCCCATC 2760
Db |||||
Qy 3071 CCTTGCCCTGCTCCAAAGGSCACAGCGAGTGCCCGAGCGGAGAGCTGATCCCCATC 3130
Db |||||
Qy 2761 CTGACGACCAAGTCTGCTCGTCCACCCCTGCACTGGTGTGGGCGAGTGTGGTCTTCCAGT 2820
Db |||||
Qy 3131 CTGGACGACCAAGTCTGCTCGTCCACCCCTGCACTGGTGTGGGCGAGTGTGGTCTTCCAGT 3190
Db |||||
Qy 2821 CTCAGCGGCGTGAAGACAAAGTGCACTCTGACTCTTATTAACAGGATAAAGTGTGGCAAC 2880
Db |||||
Qy 3191 CTCAGCGGCGTGAAGACAAAGTGCACTCTGACTCTTATTAACAGGATAAAGTGTGGCAAC 3250
Db |||||
Qy 2881 ATCACATTTACCTTTTAAACAGGAGATGATGTCACAGGTCTTACTACGAGGACATTTGC 2940
Db |||||
Qy 3251 ATCACATTTACCTTTTAAACAGGAGATGATGTCACAGGTCTTACTACGAGGACATTTGC 3310
Db |||||
Qy 2941 AGTGAATTTGAGGAATTTGAATATTTTGAAGAAATGTTTCCGCTGAATATTTCAATCTACATC 3000
Db |||||
Qy 3311 AGTGAATTTGAGGAATTTGAATATTTTGAAGAAATGTTTCCGCTGAATATTTCAATCTACATC 3370
Db |||||
Qy 3001 GCTTGGAGCCTTCCCTTTCAGCGAAACAAATGAATAATATATGTTGGCCATTTCTGTGAAGAT 3060
Db |||||
Qy 3371 GCTTGGAGCCTTCCCTTTCAGCGAAACAAATGAATAATATATGTTGGCCATTTCTGTGAAGAT 3430
Db |||||
Qy 3061 ATACGGGATGATGGAAACCCGATCAAGGAATCACTGACAAAATAATCGATCTTGTAGT 3120
Db |||||
Qy 3431 ATACGGGATGATGGAAACCCGATCAAGGAATCACTGACAAAATAATCGATCTTGTAGT 3490
Db |||||
Qy 3121 AAACGTGATGGAACAGCTCCTGATTTGCTGCCGTTGCAGAAATGAAGTTTCAGAGCGG 3180
Db |||||
Qy 3491 AAACGTGATGGAACAGCTCCTGATTTGCTGCCGTTGCAGAAATGAAGTTTCAGAGCGG 3550
Db |||||
Qy 3181 CCTCTGAAGAACAGAAACAGATTTCTTGTTCCTTGTGAGCTCTGTCTTAACTGTGGCT 3240
Db |||||
Qy 3551 CCTCTGAAGAACAGAAACAGATTTCTTGTTCCTTGTGAGCTCTGTCTTAACTGTGGCT 3610
Db |||||
Qy 3241 TGGATCTGTTGCTTGTGACGGCTTCTACTGTGCTGCGGAAGCGCGGAAGCGCGGC 3300
Db |||||
Qy 3611 TGGATCTGTTGCTTGTGACGGCTTCTACTGTGCTGCGGAAGCGCGGAAGCGCGGC 3670
Db |||||
Qy 3301 AGCCACACACACTCAGCCCTCTGAGGACAAACACCAACACACACACACACACACACAC 3360
Db |||||
Qy 3671 AGCCACACACACTCAGCCCTCTGAGGACAAACACCAACACACACACACACACACAC 3730
Db |||||
Qy 3361 CAGATCAAAAACCCCATTTGAGAAACATGGGGCCAAACACCGTCCCCATCAAGGATTAAG 3420
Db |||||
Qy 3731 CAGATCAAAAACCCCATTTGAGAAACATGGGGCCAAACACCGTCCCCATCAAGGATTAAG 3790
Db |||||
Qy 3421 AACAGAACTCCAAATGCTTAAATAGGACACACAACTTCTGAAGTAGAAGAGAGAGAC 3480
Db |||||
Qy 3791 AACAGAACTCCAAATGCTTAAATAGGACACACAACTTCTGAAGTAGAAGAGAGAGAC 3850
Db |||||
Qy 3481 ATGGACAAACACCCAGCAGAAAGCCGGTTTGGCAAGCAGCGCGGTTATACGCTGGTAGAC 3540
Db |||||
Qy 3851 ATGGACAAACACCCAGCAGAAAGCCGGTTTGGCAAGCAGCGCGGTTATACGCTGGTAGAC 3910
Db |||||
Qy 3541 AG 3600
Db |||||
Qy 3911 AG 3970
Db |||||
Qy 3601 GACAAACAGAGACTTGGAAAGTGGCCAGAGCTTAAACCGAATGGAGTACATCGTATAG 3657
Db |||||
Qy 3971 GACAAACAGAGACTTGGAAAGTGGCCAGAGCTTAAACCGAATGGAGTACATCGTATAG 4027
Db |||||

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 09:31:09 ; Search time 52 Seconds
(without alignments)
1974.294 Million cell updates/sec

Title: US-10-650-650-18

Perfect score: 6248

Sequence: 1 MRSPTGRGRPLSLALL.....LIAAIAEVRQRRPLKNRTD 1067

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 80.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6024.5	96.4	1220	2 A56136	jagged protein pre
2	2339.5	37.4	1408	2 S16148	gene serrate prote
3	1583.5	25.3	2524	2 A35844	Xotch protein - Af
4	1580.5	25.3	2703	1 A24420	notch protein - fr
5	1562	25.0	2471	2 A49128	cell-fate determin
6	1555	24.9	2437	2 S42612	transmembrane prot
7	1549	24.8	2531	2 A46019	notch-1 protein -
8	1538.5	24.6	2531	2 S18188	notch protein homo
9	1532	24.5	2555	2 A40043	notch protein homo
10	1498.5	24.0	2321	2 S78549	notch3 protein - h
11	1495.5	23.9	2531	2 T31070	notch homolog - se
12	1492	23.9	2352	2 T30201	notch homolog prot
13	1488	23.8	1064	2 A40136	fibropellin Ia - s
14	1481.5	23.7	2318	2 S45306	notch 3 protein -
15	1467.5	23.5	1203	2 A49175	Notch B protein -
16	1363	21.8	1964	2 T09059	notch4 - mouse
17	1348	21.6	728	2 I50719	C-Delta-1 - chicke
18	1342	21.5	833	2 S19087	gene Delta protein
19	1337	21.4	832	2 A13246	neurogenic protein
20	1337	21.4	880	2 S00670	neurogenic repetit
21	1321.5	21.2	722	2 I48324	DELTA-like 1 - mou
22	1261.5	20.2	2139	2 A35672	crumbs protein - f
23	1243	19.9	685	2 JC7570	Delta-4 protein -
24	1238.5	19.8	686	2 JC7569	Delta-4 protein -
25	1212	19.4	861	2 S06434	Notch homolog Motc
26	969.5	15.5	1429	2 A06434	homeotic protein 1
27	872	14.0	1372	2 T25933	hypothetical prote
28	826.5	13.2	1722	2 E89753	protein Flic7.4 [i
29	825.5	13.2	2918	2 A54105	fibrillin-2 precur

30	824	13.2	2907	2 A57278	fibrillin-2 precur
31	813	13.0	473	2 A56175	adhesive plaque pr
32	804.5	12.9	1295	2 A32901	glp1 protein precu
33	797.5	12.8	2871	2 A55567	fibrillin I - bovi
34	796.5	12.7	3002	2 A47221	fibrillin 1 precur
35	792.5	12.7	2871	2 A55624	fibrillin-1 precu
36	785	12.6	570	2 A48836	fibropellin C prec
37	780	12.5	1574	2 T13954	MEGF6 protein - ra
38	779	12.5	1620	2 T27283	hypothetical prote
39	756	12.1	1025	2 T42626	secreted leucine-r
40	740.5	11.9	4135	2 T42629	tenascin-X - bovin
41	739	11.8	3051	2 S42373	hypothetical prote
42	733	11.7	1523	2 T13953	MEGF5 protein - ra
43	723	11.6	3566	1 A40701	tenascin-X precurs
44	700	11.2	4006	2 T09070	probable tenascin
45	699.5	11.2	2201	2 A32160	tenascin-C - human

ALIGNMENTS

RESULT 1

A56136

jagged protein precursor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 21-Jul-2003

C;Accession: A56136

R;Lindseil, C.E.; Shawber, C.J.; Boulter, J.; Weinmaster, G.

Cell: 80-909-917-9095-

A;Title: 8099917-9095- a mammalian ligand that activates Notch1

A;Reference number: A56136; MUID:95211842; PMID:7657721

A;Accession: A56136

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1220 <LIN>

A;Cross-references: UNIPARC:UPI000017A1B7; GB:L38483

F;379-410/Domain: EGF homology <EGF1>

F;492-523/Domain: EGF homology <EGF>

F;634-665/Domain: EGF homology <EGF2>

Query Match 96.4%; Score 6024.5; DB 2; Length 1220;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 1026; Conservative 14; Mismatches 24; Indels 7; Gaps 2;

QY	1	MRSPTGRGRPLSLALLCALRAKVCASQGFLEILSMQNVANGELQNGCCGARN	60
DB	1	MRSPTGRGRPLSLALLCALRAKVCASQGFLEILSMQNVANGELQNGCCA--E	57
QY	61	PG-----DRKTRDECDTYFKVCLKEYQSRVTAGGPCSFGSGSTPVIGGTFNLKASRGND	116
DB	58	PGTLVRPYKTRDECDTYFKVCLKEYQSRVTAGGPCSFGSGSTPVIGGTFNLKASRGND	117
QY	117	RNRIVLPFSFAMPRSYTLVAEAWDSNDTVPDSIIIEKASHGMINPSRQWTLKONTGV	176
DB	118	RNRIVLPFSFAMPRSYTLVAEAWDSNDTVPDSIIIEKASHGMINPSRQWTLKONTGI	177
QY	177	AHFEYQIRVTCDYYGFGCNKFCRPRDDFFGHYACDQNGNKTCEMGWGPENCAICRQ	236
DB	178	AHFEYQIRVTCDHYGFGCNKFCRPRDDFFGHYACDQNGNKTCEMGWGPENCAICRQ	237
QY	237	GCSPKHSGKLPQDCRCQYGMWGLYCDKICIPHPGCVHVICNEPWQCLCETNWWGQLCDKD	296
DB	238	GCSPKHSGKLPQDCRCQYGMWGLYCDKICIPHPGCVHVICNEPWQCLCETNWWGQLCDKD	297
QY	297	LNVCYGHQPCINGGTCNTGPKYQCSPEGYSGPNCETAEHACTLSDPCNHRGCKETSLS	356
DB	298	LNVCYGHQPCINGGTCNTGPKYQCSPEGYSGPNCETAEHACTLSDPCNHRGCKETSLS	357
QY	357	GFECECSPGWTGPTCTSTNDDCSNNCSHGGTCQDLVNGFKVCVCPQWTKTQOLDANEC	416
DB	358	GFECECSPGWTGPTCTSTNDDCSNNCSHGGTCQDLVNGFKVCVCPQWTKTQOLDANEC	417
QY	417	EAKPCVNAKSKNLIASYYCDCLPGWGMQNCNDININDCLGQCQNDASCRDLVNGYRCICP	476


```
Db 652 CATSPCRNGGCEVDMGKFNKFCPLGYSGSLCEBAKENCTPSPCLEG-HCLNTPEGYICH 710
Qy 551 CPEDYEGKNCISHLKDHCRITTPC-----EVIDSCTVAMASNDTPGVRYYI---SSNYC 599
Db 711 CPDRAKHCQLRPLCSQPPNCGCFANVSLATSAITTTTTTTTATTRKMAKPSGLPC 770
Qy 600 GPHCKKCSQSGGKFTCDNKGFTTYCHENINDCESNPNCRNGGTCIDGVNSYKICISDGM 659
Db 771 SGHGSCEMSDVGTF-KCHVGHGTGTFCEHNLNECSNPNCRNGGICLDG-----DG- 819
Qy 660 EGAYCETNINDSQNPCHNGGTCRDVLNDFYCDCKNGWKGTCHSRDSQDEATCNGGT 719
Db 820 -----MPG-----APDKALQPH-----CRCAPGWTGLFC 876
Qy 720 CYDEGDAFKCMCPGMEGTTCTNIARNSSCLPNPCHNGGTCVNGESFTVCYCKEWEGPIC 779
Db 852 C-----MPG-----APDKALQPH-----CRCAPGWTGLFC 876
Qy 780 AONTNDCSPHPCVNSGTCTVDGDWYRCCECAPGAFAGPDCRININECOSPCAFGATCYDEI 839
Db 877 AEAIQCRGQFCHNGGTCESGAGFRVCAGQFSGPDCRINVNECSPPQCGGATCIDGI 936
Qy 840 NGYRCVCPGHSAGKCEVSRP-----CITMGSVIPDGAKM-----DDD 879
Db 937 GGVSCICPPRHHGLURCBILLSDPKSACQNASNTISPALTNRSONWLIDIALTGRTEDDEN 996
Qy 880 CNTCQCLNGRIAGSKVWCGRPCL-----LHKHGS---ECPSGQSCIPILDQCFVHPCTG 932
Db 997 CNACVCENGTSRCTLWGLPNCYKVDPLSKSSNLGVCKQHEVCVFPALSETCLSSFCNV 1056
Qy 933 VGECRSSSLQ-----PVTKC-TSDSYQDNCAINITFENKEMSPGLTTEHCSEL 983
Db 1057 RGDRALEPERRVAPPLPAKSSCPWQNAVVENCARLTILLALERVKGASVEGLCSLV 1116
Qy 984 RNL---NILKNVSAEYS-----IYIACEPSPANNEIHWASIEDIRDDGNPKIETDKI 1035
Db 1117 RVLLAAQLIKPASTFQDQFGLMVLCLDGTGNTDTVELTVSSSKLMDPOLPV-AVGLL 1174
Qy 1036 IDLVSRKRDGNSLLIAAFAVRVQR 1060
Db 1175 GELLSSRLNG-----IQRR 1189

RESULT 3
A35844
Xotch protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 12-Oct-1990 #sequence_revision 12-Oct-1990 #text_change 05-Oct-2004
R:Coffman, C.; Harris, W.; Kintner, C.
S:Science 249, 1438-1441, 1990
A:Title: Xotch, the Xenopus homolog of Drosophila notch.
A:Reference number: A35844; MUID:90385285; PMID:2402639
A:Accession: A35844
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: mRNA
A:Residues: 1-2524 <COF>
A:Cross-references: UNIPARC:UPI000004F253
C:Superfamily: notch protein; ankyrin repeat homology; EGF homology
C:Keywords: transmembrane protein
F:146-177/Domain: EGF homology <EGX1>
F:184-215/Domain: EGF homology <EGF1>
F:222-254/Domain: EGF homology <EGF>
F:456-487/Domain: EGF homology <EGX2>
F:757-788/Domain: EGF homology <EGF3>
F:1025-1056/Domain: EGF homology <EGX3>
F:1924-1956/Domain: ankyrin repeat homology <AN1>
F:1957-1989/Domain: ankyrin repeat homology <AN2>
F:1991-2023/Domain: ankyrin repeat homology <AN3>
F:2024-2056/Domain: ankyrin repeat homology <AN4>
F:2057-2089/Domain: ankyrin repeat homology <AN5>
```

Query Match

25.3%; Score 1583.5; DB 2; Length 2524;

```
Best Local Similarity 30.7%; Pred. No. 1.9e-83;
Matches 348; Conservative 126; Mismatches 335; Indels 323; Gaps 40;

Qy 62 GDRKTRDECDTVFKVCKEYQSRVTAGG-----PC-----SFG 95
Db 19 GLAPCTQTA-----EMCLNGRCCEMFGGTVCLGCLNLYFGERCQFPNPTCTIKQCMNFG 73
Qy 96 SGSTPVIGMFTNLKASRGNDRNRIVLFPFANPRSYTLVLEAWSSNDTVQPDISIIEKA 155
Db 74 T-CEPVLQGN-----AIDFICHCPVGFT-----DKV-----CLTPV 103
Qy 156 SHSGMNLPSQWOTLQNTGVAHPEYQIRVTCDDYYFGCNKF-----CRP 202
Db 104 DNACVANNPCRNCGTCELLNSVT--EYKCR--CPPGMTGDSQQADPCASNPANGGKCLP 159
Qy 203 RDPFFGHYACDQNGKNTCMEGWMP-----ECNRAICRQG--CSPKHGSKCLPGDCRC 253
Db 160 ---PEIQYIC-----KCPGFHGAATCKQDINESQNPKNGGQCINEFQSYR---CTC 206
Qy 254 QYQWQGLYCDK--CIPHPGCVHGI CNE-----PMWCLCTNNWGGQLCDKDLNYCGTHQP 305
Db 207 QNRFTGRNCDPEYVVPNPSPLNGGTCRQTDTSYDCTCLPGFSGQNCENIDDCPSNN- 265
Qy 306 CLNGGTCSTNTPDKYOCSCPEGYSGPNC--EIAEHAACLSDPCHNRGSKETSIGFECECS 363
Db 266 CRNGGTCVD--GVNTYNCQCPDWTGOYCTEDVDECQMLPNACQNGGTCHNTYGGYNCV 324
Qy 364 PGWTGPTCSNIDDCSPNNCSHGTCQDLVNGFK-----397
Db 325 NGWTGEDSENIDDCANAAHSGATCHDRVASFYCEPCPHORTGLLCHLDNACISNPNCEG 384
Qy 398 -----CVCPPQWTKTQOLDANECE--AKPCVNAKSKCNLIAYYCDCLPGW 442
Db 385 SNCDTNPVNGKAITCTPGYTGPACNNDVDECSILGANPCERGCRTNLTGSLFCQNCPOG 444
Qy 443 MGQNCIDINIDCLGO--CONDASCRDLVNGYRCICPGYAGDHCDERDIDECASNPCLNGH 501
Db 445 AGPRCIDVNECLSNPCQNDSTCLDQIGEFQCICMPGYEGLYCETNIDECASNPCLNGK 504
Qy 502 COMEINRFQCLPTGFSGNLCQDIDYCEPNPCQNGAQCYNRASDYCECKPEDEYEGKCS 561
Db 505 CIDKINEFRCDPTGFSGNLCQHDDECTSTPCNKAQKCLDGPNSYTCQCTEGFTGRHCE 564
Qy 562 HLKDHCRITTPCEVIDSC-----TVAMASND-----586
Db 565 QDINECIPDPCH-YGTCCKDGIATFTCLCRPGYTGRLCDNDINECLSKPCLNGGQCTDREN 623
Qy 587 -----TPEGVRYI-----SSNVCGPHGKCKSSGGKFTCDCKNGFTGTTCHEINID 632
Db 624 GYICTCPKGTGTGVNCTETKIDDCASNLCDNGKCIDKIDG-YECTCEPGYTGKLCNININE 681
Qy 633 CESNPCRNGGT-----CIDGVNSYKIC 655
Db 682 CDSNPCRNGTCKDQINGFTCTVCPDGYDHDMCLSEVNECSNFCIHGACHDGVNGYKDC 741
Qy 656 SDGMEGAYCETNINDSQNPCHNGGTCRDVLNDFYCDCKNGWKGTCHSRDSQDEATCN 715
Db 742 EAGWSGNCIDINNNECESNPNNGGTCXDMTGAVICTCKAGFSGPNCQTNINESSNPCL 801
Qy 716 NGTCTYD-----EGDAFKCMCPGGW 735
Db 802 NHGTCIDDVAGYKCNMLPYTGAICEAVLAPCAGSPCKNGGRCKESDFTFSCPCPGW 861
Qy 736 EGTTCNIARNSSCLPNPCHNGGTCVNVGESFTVCYCKEWEGPICCAONTNDCSPHPCVNSG 795
Db 862 QGQTCIDWN--ECVNRPCRNGATQNTNGSKNCKPGYTGRCNCEWIDIDCQNPCHNGG 920
Qy 796 TCVDGDNWYRCCECAPGAFAGPDCRININECOSPCAFGATCVDREINGYRCVCPFGHSGAK 855
Db 921 SCSDGINMPFCNCFAGFRGPKCEEDINECASNFKNGANCTDCVNSTCTCTCQFGFSGIHC 980
Qy 856 QEVSGRPCTIMGSVIPDGAKWDDDCNTCQCLNRIACSKVWCGRPCL--LHKHGSCEP- 912
```


Db 981 E--SNTF-----DCTESSCFNG-----GTCIDGINTFTCCQPP 1011

Qy 913 --SQSCIPILDQCFVHPCTGVGECRSSLSLPVKTKTSDSYQD-NCANI 961

Db 1012 GFTGSYQCHDI-NECDSKPCLNGTCDQSY---GTYKCTCPQGYTLNQNQL 1059

RESULT 4

A24420

notch protein - fruit fly (*Drosophila melanogaster*)

N:Alternate names: neurogenic repetitive locus protein

C:Species: *Drosophila melanogaster*

C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 05-Oct-2004

C:Accession: A24420; A24768; S09358; A05267

R:Kidd, S.; Kelley, M.R.; Young, M.W.

Mol. Cell. Biol. 6, 3094-3108, 1986

A:Reference number: A24420; MUID:87064624; PMID:3097517

A:Accession: A24420

A:Molecule type: DNA

A:Residues: 1-2703 <KID>

A:Cross-references: UNIPROT:P07207; UNIPARC:UPI000016BCC6; GB:K03508; NID:gl57991; PIDN:

R:Wharton, K.A.; Johansen, K.M.; Xu, T.; Artavanis-Tsakonas, S.

Cell 43, 567-581, 1985

A:Reference number: A24768; MUID:86079539; PMID:3935325

A:Accession: A24768

A:Molecule type: mRNA

A:Residues: 1-48, 'I', 50-118, 'R', 120-230, 'I', 232-256, 'N', 258-266, 'A', 268-872, 'R', 874-958, 'R', 959-1012

A:Cross-references: UNIPARC:UPI0000173D1F

A:Note: the authors translated the codon ATC for residue 49 as Thr, ATT for residue 2044

R:Tautz, D.

Nucleic Acids Res. 17, 6463-6471, 1989

A:Title: Hypervariability of simple sequences as a general source for polymorphic DNA

A:Reference number: S09358; MUID:89385974; PMID:2780284

A:Accession: S09358

A:Molecule type: DNA

A:Residues: 2505-2551, 'QQQQ', 2552-2576, 'E', 2578-2604 <TAU>

A:Cross-references: UNIPARC:UPI0000173D20

R:Wharton, K.A.; Vedyobnick, B.; Finnerty, V.G.; Artavanis-Tsakonas, S.

Cell 40, 55-62, 1985

A:Title: opa: a novel family of transcribed repeats shared by the Notch locus and other

A:Reference number: A05267; MUID:85099329; PMID:2981631

A:Accession: A05267

A:Molecule type: DNA

A:Residues: 2504-2576, 'E', 2578-2611 <WHA2>

A:Cross-references: UNIPARC:UPI0000173D21

C:Genetics:

A:Gene: notch; opa

A:Cross-references: Flybase:FBgn0004647

A:Map position: 8.96-9.36

A:Introns: 53/3; 84/3; 171/3; 240/3; 283/3; 2333/3; 2436/3; 2588/3

C:Superfamily: notch protein; ankyrin repeat homology; EGF homology

C:Keywords: differentiation; tandem repeat; transmembrane protein

F:27-43/Domain: transmembrane #status predicted <TMW1>

F:297-328/Domain: EGF homology <EGX1>

F:530-561/Domain: EGF homology <EGF1>

F:568-599/Domain: EGF homology <EGF>

F:988-1019/Domain: EGF homology <EGX2>

F:1064-1095/Domain: EGF homology <EGF3>

F:1187-1218/Domain: EGF homology <EGX3>

F:1746-1762/Domain: transmembrane #status predicted <TMM2>

F:1950-1982/Domain: ankyrin repeat homology <AN1>

F:1983-2015/Domain: ankyrin repeat homology <AN2>

F:1988-2004/Domain: transmembrane #status predicted <TMM3>

F:2017-2049/Domain: ankyrin repeat homology <AN3>

F:2050-2082/Domain: ankyrin repeat homology <AN4>

F:2083-2115/Domain: ankyrin repeat homology <AN5>

F:2538-2568/Region: glutamine-rich

F:2538-2568/Domain: neurogenic repetitive element #status predicted <OPA>

Query Match 25.3%, Score 1580.5; DB 1; Length 2703;

Best Local Similarity 28.1%; Pred. No. 3e-83;

Matches 374; Conservative 139; Mismatches 432; Indels 385; Gaps 47;

```
Db 1219 ELNIDDCAPNCPQNGTCHDRVMNFSCSPGWTGIIICBINKODCKPGACHNNGSCIDRV 1278
Qy 859 SGRPCITWGSVIPDGAWDDDCNTC---QCLN-GRIAC---SKVMCGPRPCLLHKHSE 910
Db 1279 GGFECVCPQGFV--GARCEGDINECLNPNCSNAGTLCVQLVNNYHCNCRP-----GH-- 1329
Qy 911 CPSSQSCIPILDQCFVHPCTGVGEGR-----SSSLQPVTKTSDSYQ 955
Db 1330 --MGRHCEHKV-DFCAQSPQNGGNCNIRSGHHCICNNGFYGNCKNCLSGQDCDSNCRV 1386
Qy 956 DNCANITFTNKEMMSP-GLTTEHCSELNRLNLKNVSEYSYIACESPSPANNEIHV 1014
Db 1387 GNCVVADEGFGYRCEPRGTGHE-----CEIDTLDECSNPACQ----- 1426
Qy 1015 AISAEIRD 1024
Db 1427 GAACEDLLGD 1436

RESULT 5
A49128
cell-fate determining gene Notch2 protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 05-Oct-2004
C;Accession: A49128
R;Weinmaster, G.; Roberts, V.J.; Lemke, G.
Development 116, 931-941, 1992
A;Title: Notch2: a second mammalian Notch gene.
A;Reference number: A49128; MUID:93202015; PMID:1295745
A;Accession: A49128
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-2471 <WE1>
A;Cross-references: UNIPROT:Q9QW30; UNIPARC:UPI000004F255
A;Experimental source: Schwann cell
A;Note: Sequence extracted from NCBI backbone (NCBIP:127811)
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology
F;264-295/Domain: EGF homology <EGX1>
F;799-830/Domain: EGF homology <EGF1>
F;877-908/Domain: EGF homology <EGX2>
F;1029-1060/Domain: EGF homology <EGF>
F;1067-1098/Domain: EGF homology <EGX3>
F;1153-1184/Domain: EGF homology <EGF3>
F;1191-1222/Domain: EGF homology <EGX4>
F;1876-1908/Domain: ankyrin repeat homology <AN1>
F;1909-1941/Domain: ankyrin repeat homology <AN2>
F;1943-1975/Domain: ankyrin repeat homology <AN3>
F;1976-2008/Domain: ankyrin repeat homology <AN4>
F;2009-2041/Domain: ankyrin repeat homology <AN5>

Query Match 25.08; Score 1562; DB 2; Length 2471;
Best Local Similarity 29.68; Pred. No. 3.3e-82;
Matches 326; Conservative 122; Mismatches 326; Indels 328; Gaps 34;

Qy 186 TCDDYVYFGCNKFCRRDDFFGHYA-----CDQN-----GNKT--CWEGW 224
Db 40 TCVTYHNGTG---YCRPEGFLGEYCOHRDPCEKNRCQNGGTCVTQAMLKATCRCAPGF 96
Qy 225 MGPECNRAI-----CRQGSCKHGSCKLPG-----DCRCQVWQGLYC-----DKCIPH 268
Db 97 TGSDCQYSTHPCFVSFRPCQNG-----GTCHLSWDTYECTCQVGFSGKOCOWTDVCLSH 151
Qy 269 PGCVHVG-----ICNEPWOCLCETNWWGQCLDKLNYCGTHQPCLANGTCSNTGPDKYQCS 323
Db 152 P-CENGTSCSVANQ--FSCRCAPAGITQKCDADINECDIPRCQHQGTCLNL--PGSTRQ 208
Qy 324 CPFGYSGNCEIAEHACLDPCHNRGSKET--SLGFCEGSPGWGTCTSTNIDDCSPNN 382
Db 209 CPQRFTHQCHDSPPVPCAPSPVNGGTCRGTGDTSECHCLPGFEGSNCRNIDDCPNHK 268
Qy 383 CSHGCTQDILVNGFKVCVCPQWTGKTCQLDANEC----- 416
```

```
Db 269 CQNGGVCVGVNTYNCRCPPQWTGQFCTEDVDBCLLOPNACQNGGTCNRRNGGVCVCVN 328
Qy 417 ----- 416
Db 329 GWSGDDCSENIDDCAFASCTPGSTCIDRVASFCLCPGKAGLLCHLDDACISNPKHGA 388
Qy 417 -----EAKPCVNAKSKRLIAYVYCDCLPGW 442
Db 389 LCDTNPLNGQYICTCPQAYKAGDCTEDVDECAMANSNPCEHAGKCVNTDGAFHCECLGY 448
Qy 443 MGQNCNININDLQO--CONDASCRDLVNGYRCICPPGYAGDHCBERRDIDECASNPCLNGGH 501
Db 449 AGPRCEMDINECHSDPCQNDATCLDKIGFTCLCMFGFKGVHCELVNEVQSPCVNNGQ 508
Qy 502 CQNEINRFQCLCTPGFSGNLQCLDIDYCEPNPCQGAQCYNRASDYFCCKPEDEYEGNCS 561
Db 509 CVDKVNRFQCLCPPGFTGPVCQIDIDDSCSTPLNGAKCIDHPNGYECQCATGFTGLCD 568
Qy 562 HLKDHCRTPC-----EVIDSCTVAMASND----- 586
Db 569 ENIDNCDPDPCHRGQCCQDGIDSYTCICNPGYMAICSDQIDECYSFPCLDNGRCIDLWNG 628
Qy 587 -----TPGVRY-----ISSNVCGPHGCKKSQSGKFTCDCKNGFTCTYCHENINDC 633
Db 629 YQCNCPQPGSLGNCENIFDDCASNPC-LHGAC-VIGINRYSVCVSPGFTGQRNIDIDEC 686
Qy 634 ESNPCRNNGTCTIDGVN-----SYKICIS 656
Db 687 ASNPCKDATCINDVNGFRCMCPGPHHPSCYSQVNECLSSPCIHGNCCTGGLSGYKCLCD 746
Qy 657 DHEGAYCETNINDSONPCHNGGTCRDLDVNDFYCDCKNGWKTKTCHSRDSQDEATCNN 716
Db 747 AGWVGINCEVDKNECLSNPCQNGGTCNNLVNGYRCTCKGFKGYNCQVNIDECASNPCLN 806
Qy 717 GCTCYDEGDAFKCMCPGWEHTTCNARNSSCLPNPCHNGGTC--VVGESFTCVCKEGW 774
Db 807 QGTCLLDDVSGYTCMCLPYTKNCQTVL-APCSNPNCENAAVCKEAPNFSFSTCLCAPGW 865
Qy 775 EGPICAQNTNDCSPHPCYNSGTCTVDGDNWYRCAPGAFGDCRININECOSSPCAFAT 834
Db 866 QGQRCCTVDVDECVSKEPMNGICHNTQGSYMCCEPFGSGMDCCEEDINDCLANPCQNGS 925
Qy 835 CVDEINGYRCVCPGHSKAKCO-----EVSGRPCITWGS-----VLPDG-----A 874
Db 926 CVDKNTVFSCLCLPGFVGDKCQTDNNECLSEPCCKNGGTCSDYVNSYTCCTCPAGFHGVCE 985
Qy 875 KWDDDCNTCCLNG-----RIACSKVMCGP-----RPCI-----LH 905
Db 986 NNIDECTESSCFNGGTCVDGINSFSCLCVPVGFPGFCLHDINECSNPCLNSGTCTVDGLG 1045
Qy 906 KGHSECP---SGQSCIPILDQCFVHPCTGVGECRSSSLQPVTKTCTS-----DSYQDNC 958
Db 1046 TYRCTCTPLGYTKNC-QTLVNLCSPPCKNKGTCQAQEKARP---RCLCPGWDGAY---C 1098
Qy 959 ANITFTFNKEMMSPGLTTEHIC 980
Db 1099 DVLNVSCAAALQKGVVVEHLC 1120

RESULT 6
S42612
transmembrane protein precursor - zebra fish
C;Species: Brachydanio rerio (zebra fish)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: S42612
R;Bierkamp, C.; Campos-Ortega, J.A.
Mech. Dev. 43, 87-100, 1993
A;Title: A zebrafish homologue of the Drosophila neurogenic gene Notch and its pattern o
A;Reference number: S42612; MUID:94128602; PMID:8297791
A;Accession: S42612
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2437 <BIE>
```

A;Cross-references: UNIPROT:P46530; UNIPARC:UPI000013056E; EMBL:X69088; NID:9433866; PDB:1W55; C;Superfamily: notch protein; ankyrin repeat homology; EGF homology
F;755-786/Domain: EGF homology <EGF1>
F;1023-1054/Domain: EGF homology <EGF>
F;1185-1216/Domain: EGF homology <EGF2>
F;1315-1947/Domain: ankyrin repeat homology <AN1>
F;1348-1980/Domain: ankyrin repeat homology <AN2>
F;1382-2014/Domain: ankyrin repeat homology <AN3>
F;2015-2047/Domain: ankyrin repeat homology <AN4>
F;2048-2080/Domain: ankyrin repeat homology <AN5>

Query Match 24.9%; Score 1555; DB 2; Length 2437;
Best Local Similarity 32.8%; Pred. No. 8.1e-82;
Matches 321; Conservative 115; Mismatches 317; Indels 225; Gaps 34;

QY	200	CRPDDFFGHYACDQNGKNT-----CMQEW-----MGPECNRAICROGCSPKHGSK	246
DB	73	CRPQ-----MQNEVGKCDCLVGFSDRLCLTPVNHACMNSPCRG-----GTCS	117
QY	247	L-----PGDCRCQYQWGLYC---DKCIPHCPCVHGICN---EPMQCLCETNWGQLCDX	296
DB	118	LLTLDTFTTCRCPCGWSGKTCQLADPCASNPCCANGGQCSAFESHVICTCPNPFHGTQRQD	177
QY	297	LYNYCG--THQPCINGGTSNTGPKDYKQCSCEGYSGNPCEIAEHACLSDPCNHRGSKETS	355
DB	178	VNECAVSFSPCRNGGTCINE--VGSYLRCRCPPEYTPHCQRLLYQCLPSPCRSGGTCVQTS	236
QY	356	-LGFECCSPGWTGPTCTSTNIDDCSPNNCSHGCTQDLVNGFKVCVPPQWTKTKCQLDAN	414
DB	237	DTHTTCSCLPGFTGQTCEHNVDCTQACENGGPCIDGINTYNCHDKHWTGQVCTDEVD	296
QY	415	ECEAKP--CVNAKSKNLIASYDCPLPGWGMQNCININDCL--GQCQNDASCRDLVNGY	471
DB	297	ECELSPNACQNGGTCNTIGGFHCVCVNGWTGDDCSENIDDCASACSHGATCHDRVASF	356
QY	472	-----RCIPPGVAGDHCEIDIIECA	492
DB	357	FCECPHGRTGLLCHLDDACISNQCQGSNCDTNPSGKAICTCPPGYTGSACNQDIDEC	416
QY	493	--SNPCLNGHCQNEINRFQCLPTGPSNLCLQDIDYCEPNPCQNGAOCVNRASDYFCK	550
DB	417	LGANPEHGGCLNLTGKSGFCKCLQYEGPRCEMDVNECKSNPCQNDATCLDQIGGFHCI	476
QY	551	CPEDYEGKNCSHLKHDCRTTPECEVIDSCVTAMASNDTPEGVYISSNVGPHGKCKSQSG	610
DB	477	CMPGVEGVFCQINSDDCASQFC-----LNGKCIDKI--	507
QY	611	GKFTCDNCNGTGYTYCHENINDCSNPNCRNGGTCIDGVNSYKICISDGWEGAYCETIND	670
DB	508	NSFHECEPKGFGSGLCQVDVDECASTPCKNGAKCTDGPENKYTCBCTPGFSIHGELDINE	567
QY	671	CSQNPCHNGGTCRLVNDIFYCDCKNGWKTKCHSRDSQCEATCNNGTCVDEGDAPKCM	730
DB	568	CASFPCHY--GYCRDGVASFCTDCRPGYTGRICETNINECLSQPCRNNGTCDRENAYICT	626
QY	731	CP-----GGWEGTTCNIARNSSCLNPPC	753
DB	627	CPKGTGTGVCENINDDKRKPDCYKCIDKXINGVECVCEPGYSGSMCN--NIDDCALNPPC	685
QY	754	HNGGTCVNVGSPFTVCKEGWGEPFCAQNTNDCSPHPCYNSTGTCVDGNWTRCBACPGFA	813
DB	686	HNGGTCIDGVNSFTCLCPDGPFRDATCLSHQNECSSNPCIH--GSLDQINSYRCVCEAGWM	744
QY	814	GPDCEININECOSSPCATGATCVDENGYRCVCPFGHSGAKCQ----EVSGRPCITWGSV	869
DB	745	GRNCDININECLSNPCVNGGTCQMTSGYLCTCRAGFSGPNCOMNINECASNPCLNQSC	804
QY	870	IPDGAkWDDCNTCQCLNGRIACSKVM--CGPRPCL-----LKHGSECPSS--G	914
DB	805	IDDVAGF--KNCNMLPYTTEV--CENVLAPCSRPCNGGVCRESEDFQSPCNCNPAHQG	861
QY	915	QSCIPILDQCQFVHPCTGVGBCRSLSQFVKTKTSDSYQDNCANITFTTFNKEMWSPGL	974

862 QTC-EVDINECVRPCTNGVGCENLR-----GGFQCR-----NGF 897

975 TTEHICSELRLNLKVNLSAEYSIYIACEPPSPANNEI-----HVAISAEIRDGNP 1027

898 TGA-LCE-----NDIDD-----CEPNPCSGGVQCDRVNGFVVCVCLAGFR--GER 939

1028 IKEITDKIIDLVSXRDGN 1045

940 CAEDIDECVSAPCRNGN 957

RESULT 7

A46019

notch-1 protein - mouse

N;Alternate names: notch protein

C;Species: Mus musculus (house mouse)

C;Date: 22-Sep-1993 #sequence revision 18-Nov-1994 #text change 05-Oct-2004

C;Accession: A46019; S25144; C49175; B46438; A46438; PH1569; S32109

R;del Amo, F.F.; Gendron-Maguire, M.; Swiatek, P.J.; Jenkins, N.A.; Copeland, N.G.; Grid Genomics 15, 259-264, 1993

A;Title: Cloning, analysis, and chromosomal localization of Notch-1, a mouse homolog of A;Reference number: A46019; MUID:93194170; PMID:8449489

A;Accession: A46019

A;Status: not compared with conceptual translation

A;Molecule type: nucleic acid

A;Residues: 1-2531

A;Cross-references: UNIPROT:Q01705; UNIPARC:UPI000002922B; GB:S47228; NID:928

A;Note: sequence extracted from NCBI backbone (NCBIP:127318)

R;Franco del Amo, F.; Smith, D.E.; Swiatek, P.J.; Gendron-Maguire, M.; Greenspan, R.J.; I submitted to the EMBL Data Library, April 1992

A;Description: Expression pattern of Notch, a mouse homolog of Drosophila Notch, suggest A;Reference number: S25144

A;Accession: S25144

A;Molecule type: mRNA

A;Residues: 1551-2108, 'Q', 2110-2114, 'ALP', 2118-2170 <PRA>

A;Cross-references: UNIPARC:UPI0000177461; EMBL:Z11886

R;Lardelli, M.; Lendahl, U.

Exp. Cell Res. 204, 364-372, 1993

A;Title: Notch A and Notch B--two mouse Notch homologues coexpressed in a wide variety o A;Reference number: A49175; MUID:93178563; PMID:8440332

A;Accession: C49175

A;Status: preliminary; nucleic acid sequence not shown

A;Residues: 1161-1547 <LRA>

A;Cross-references: UNIPARC:UPI0000177462; EMBL:X68278; NID:9287987; PIDN:CAA48339.1; PI A;Experimental source: embryo

A;Note: sequence extracted from NCBI backbone (NCBIP:126159)

R;Kopan, R.; Weintraub, H.

J. Cell Biol. 121, 631-641, 1993

A;Title: Mouse notch: expression in hair follicles correlates with cell fate determinati A;Reference number: A46438; MUID:93252998; PMID:8486742

A;Accession: B46438

A;Status: preliminary

A;Molecule type: nucleic acid

A;Residues: 1865-1932, 'RR', 1935-1937, 'L', 1938-1967, 'I', 1969-2044, 'IE', 2047-2052, 'S', 2054 A;Cross-references: UNIPARC:UPI0000177463

A;Experimental source: embryo

A;Note: sequence extracted from NCBI backbone (NCBIN:131246, NCBIP:131247)

C;Comment: This protein has many EGF repeats and lin-12[1172]/Notch repeats.

C;Comment: This protein is one of the neurogenic proteins controlling the decision between C;Genetics:

A;Gene: notch-1

A;Map position: 2

A;Note: proximal region of chromosome 2

C;Superfamily: notch protein; ankyrin repeat homology; EGF homology

F;106-138/Domain: EGF homology <EGF1>

F;144-175/Domain: EGF homology <EGF1>

F;222-254/Domain: EGF homology <EGF2>

F;261-292/Domain: EGF homology <EGF2>

F;339-370/Domain: EGF homology <EGF3>

F;416-449/Domain: EGF homology <EGF3>

F;456-487/Domain: EGF homology <EGF4>

F;494-525/Domain: EGF homology <EGF4>

F;532-563/Domain: EGF homology <EG06>
F;607-638/Domain: EGF homology <EG07>
F;682-713/Domain: EGF homology <EG08>
F;757-788/Domain: EGF homology <EG09>
F;795-826/Domain: EGF homology <EG10>
F;873-904/Domain: EGF homology <EG11>
F;911-942/Domain: EGF homology <EG12>
F;949-980/Domain: EGF homology <EG13>
F;987-1018/Domain: EGF homology <EG14>
F;1025-1056/Domain: EGF homology <EG15>
F;1063-1094/Domain: EGF homology <EG16>
F;1149-1180/Domain: EGF homology <EG17>
F;1187-1218/Domain: EGF homology <EG18>
F;1233-1264/Domain: EGF homology <EG19>
F;1352-1383/Domain: EGF homology <EG20>
F;1391-1425/Domain: EGF homology <EG21>
F;1499-1981/Domain: ankyrin repeat homology <AN1>
F;1983-2015/Domain: ankyrin repeat homology <AN2>
F;2016-2048/Domain: ankyrin repeat homology <AN3>
F;2049-2081/Domain: ankyrin repeat homology <AN5>

Query Match 24.8%; Score 1549; DB 2; Length 2531;
Best Local Similarity 30.4%; Pred. No. 1.9e-81;
Matches 325; Conservative 132; Mismatches 357; Indels 256; Gaps 32;

Qy 91 PCFSGSGSTPVIGNTFNLKASGRNDRNRIVLFPFAMPRSYTLVBEAWDSSNDVTQVDS 150
Db 380 PCNEGS-----NCDTNPVNGKRICTCPSGYTGPA-----CSQDVDECDL 418

Qy 151 IIEKASHG-MINP-SRQWTLKQNTGVAFHFQIR-----VTCDYYGPGC- 196
Db 419 GANRCEHAGCLNTLGSFECQCLQYTGPG-CEIDVNECISNPQNDAATCLDQIGEPQCI 477

Qy 197 -----NKFCRPRDDFFGHVACDQNGN-----KTCMEGWMGPECNRAICRQCGSP- 240
Db 478 CMFGYEGYCEINTDECASSPLNGHCHMDKIHFQCQCPKGFNGHLQCQDVDECASTPC 537

Qy 241 KHGSKLPG-----DCRCQYGMQGLYC-----DKCIPHPGCVHGCNE---PWOCLCETNMG 289
Db 538 KNGAKCLDGNNTYTCVTEGYTGTHCEVDIDECDDPDP-CHYGSKDQGVAFFTCLCQPGYT 596

Qy 290 QQLCDKDLNCGTHQPCPLNGGTCSTGPDYKQSCPEGYSGPNCETAEHACLSDPCPNRG 349
Db 597 GHICETNINECHS-QPCRHGCTQCDR-DNSYLCLCLKGTGTGPNCEINLDDCASNP-DSG 653

Qy 350 SKETSIGPECECSPGWGTPTCSNIDDCSPNNCSHGTCQDLVNGFKVCVCPQWTKTC 409
Db 654 TCLKIDGYECACEPGYTGSMCNWNIDECAGSPCHNGGTCTCEDGIAGTCTCPGHDPTC 713

Qy 410 QLDANECEAKPCVNAKSKNLIASYCDLPGWMQNCIDINDC-LGQCONDASCRDLV 468
Db 714 LSEVNECNSNPCIHG-ACRDGLNGYKCDCAPGWSGTNCDINNECESNPCVNGTCKDWT 772

Qy 469 NGYRCICPPGYAGDHCBRDIDECASNPLNGHCHQNEI----- 506
Db 773 SGVYCTCREGSPNCOTNINECASNPCLNQGTCIDDVAGYKNCPLPYTGATCEVVLAP 832

Qy 507 -----NRFQCLCPTGFSGNLCQDIDICEPNPCQNGAQCVNRSADYF 548
Db 833 CATSPKNSGVKESDYBSFCVPTGWMGQGTCEVDINECVKSPCRHGASQNTGYSR 892

Qy 549 CKGPEDYEGNCSHLKHCHRTTTPCEVIDSCT--VAMASND-----TPEGVRISSN 597
Db 893 CLCQAGYTGNCESDIDDCPNPCHNGSGCTDGINAFCDCLPGFQGAFCCEEDINECASN 952

Qy 598 VCPGPKCKSQQSGKFTCDCKNGFTGYCHENINDCESNCRNGGTICIDGNSYKICSD 657
Db 953 PCQNGANC-TDCVDSYTCPCPVGFNGIHCENNTPDCTESSCFNGGTCVDGINSFTCLCPP 1011

Qy 658 GWEGAYCETNINDCSQNPCHNGTCDRLVNDVFCDCKNGHKGKTCNRSQSCDEATCNG 717
Db 1012 GFTGSYCOVDVNECDNRPCPLHGTCQDSYGTYKCTCPQGYTGLNQNQLVRWCDSPCKNG 1071

718 GTCYDEGDAFKCMCPGSGWEGTTCNIARNS----- 746
1072 GRCWQNTNQVHCRCRSGWTGVNCDVLSVSEVAARKGIDVTLCCQHGGLCVDGDKHYC 1131

747 -----SCLPNPCHNGGTCTVNGESFTVCYCKEGWEGFICAQNTNDSCSPHP 790
1132 HCQAGYTGSGYCEDEVDSCSPNCPQNGATCTDYLGFSCKCVAGYHSGNSCEEINECLSQP 1191

791 CYNSTGTCVDGDNVRCBCEPAGFAGPDCRININEC-----OSSPACAFGATCVDEINGY 842
1192 CQNGTCTIDLTNSYKSCPRGTQGVHCEINVDCHPLDPASRSPKCFNNGTCTVDQGGY 1251

843 RCVCPP-----GHSAGAKCOEV----- 858
1252 TCTCPGFGVGERCEGDVNECLSNPCDPRTQNCVQRVNDPFHCECRAGHTGRRCESVINGC 1311

859 SGRPCITMGSVIP-----DGAKWDDCWTC-----OCLNGRIASKWCWGP 899
1312 RGKPCNKGVCVAVASNTARFICRCFAGFEGFENDARTCGSLRCLNGTCTIS-----GP 1367

900 RP-----CLLKHGSEC--PSGQSCIPILDQCFVHPCTGVGECRSSSLOP 943
1368 RSPTCLCLSGSFTQCEQFPASSPCVG-----SNPCYNQGTCEPTSENP 1410

RESULT 8
S18188
notch protein homolog - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 02-Aug-2002
C;Accession: S18188
R;Weinmaster, G.; Roberts, V.J.; Lemke, G.
Development 113, 199-205, 1991
A;Title: A homolog of Drosophila Notch expressed during mammalian development.
A;Reference number: S18188; MUID:92111383; PMID:1764995
A;Accession: S18188
A;Molecule type: mRNA
A;Residues: 1-2531 <WEI>
A;Cross-references: UNIPARC:UPI000177456; EMBL:X57405; NID:957634; PID:957635
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology
F;987-1018/Domain: EGF homology <EGF1>
F;1025-1056/Domain: EGF homology <EGF2>
F;1233-1264/Domain: EGF homology <EGF2>
F;1917-1949/Domain: ankyrin repeat homology <AN1>
F;1950-1982/Domain: ankyrin repeat homology <AN2>
F;1984-2016/Domain: ankyrin repeat homology <AN3>
F;2017-2049/Domain: ankyrin repeat homology <AN4>
F;2050-2082/Domain: ankyrin repeat homology <AN5>

Query Match 24.6%; Score 1538.5; DB 2; Length 2531;
Best Local Similarity 31.7%; Pred. No. 7.5e-81;
Matches 305; Conservative 120; Mismatches 316; Indels 221; Gaps 26;

Qy 186 TCDDYYGFGC-----NKFCRPRDDFFGHVACDQNGNKT-----CMEGWMGPEC 229
Db 466 TCLKIDGYEQICMPGEGYCEINTDECASSPLNGHCHQNEI----- 525

Qy 230 NRAICRQCGSP-KHGSCKLPG-----DCRCQYGMQGLYC-----DKCIPHPGCVHGCNE-- 278
Db 526 QYDVDECATPCNPKNGAKCLDGNNTYTCVTEGYTGTHCEVDIDECDDPDP-CHTGLCKDGV 584

Qy 279 -PWOCLCETNMGQCLDKDLNCGTHQPCPLNGGTCSTGPDYKQSCPEGYSGPNCETAE 337
Db 585 AFTTCLCQPGYTGHHCHETNINECHS-QPCRHGCTQCDR-DNYYLCLCLKGTGTGPNCEINL 642

Qy 338 HACLSDPCPNRGSKETSIGPECECSPGWGTPTCSNIDDCSPNNCSHGTCQDLVNGFK 397
Db 643 DDCASNPC-DSGTCLDKIDGYECACEPGYTGSMCNWNIDECAGSPCHNGGTCTCEDGIAGT 701

Qy 398 CVCPPQWTKTCDLDANECEAKPCVNAKSKNLIASYCDLPGWMQNCIDINDC-LG 456
Db 702 CRCPGHDPTCLSEVNECNSNPCIHG-ACRDGLNGYKCDCAPGWSGTNCDINNECESN 760

QY 457 QCONASCRDLVNGYRCICPPYAGDHCDERDIDECASNPCLNGHHCQNEI----- 506
Db 761 PCVNGGTCCKDWTSGYVCTCREGFGNPNQTNINECASNPCLNQGTCLDDVAGYKNCPLP 820
QY 507 -----NRFOCLCPTGFGSNLQCLDIDYCEPNPCQN 536
Db 821 YTGATCEVLAPCATSPCKNSGVCKESEDYESFSCVCTQWQGTCEIDINECVKSPCRH 880
QY 537 GAQCYNRASDYFKCPBEDYEGKNCSHLKHCRTPPCBEVDSCT--VMAASND----- 586
Db 881 GASCONTNGSYRCLCAQYTGGRNCESDIDDCRPNPCHNGGSCTDGVNAAPCDCLPFGQA 940
QY 587 -TPGVRYISNVCGPHGCKKSQSGGKFTCDCKNGFTGTTCHEINDCESNPNCRNGTCTI 645
Db 941 FCEBIDINECATNPQNGANC--TDCVDSYITCTCTPGFNGIHCENNTPDCTESSCFNGGTCV 999
QY 646 DGVNSYKICISDGMWEGAYCETNINDSQNPCHNGGTCDRLVNFYCDCKNGWKGTCHSR 705
Db 1000 DGINSFTCLCPPGFTGSGYCYQDVNECDSPCLHGGTCQDSYGYKCTCPQGYTGLNCQNL 1059
QY 706 DSQDEATONNGGTCYDEGDAFKCMCPGPGWEGTTCTIARNS----- 746
Db 1060 VRWCDSAPCKNGGCKWQTNQYHCECRSGWTGFNCVDLSVSCVAAQKRGIDVTLQCQHG 1119
QY 747 -----SCLPNPCHNGGTCVNGESFTCVCKEGWGP 778
Db 1120 GLCVDEEDKHVCHQAGYGSYCEBDEVECPNPNQNGATCTDYLGFSCKCVAGYHGSN 1179
QY 779 CAQNTNDCSPHPCVNSGTCDVDNWRCEAPGAPGAPDCRININEC-----QSSPCA 830
Db 1180 CSEINECLSQPCQNGGTCIDLTWYKCSPRGTQGVHCEINVDVDDCHPPLDPASRSPKCF 1239
QY 831 FGATCVDEINGYRCVCP-----GH 850
Db 1240 NNGTCVDQVGGYTCTCPFGVGERCEBGEVBNCLSNPCDPRGTQNCQVRVNDPFHCECRAGH 1299
QY 851 SGAKCOEV-----SGRPCLTMSVTP-----DGAKWDDDCNTC---QCLN 887
Db 1300 TGRCEVINGCRGKPCRNNGGVCAVASNTARGFTCRCPARPEGATCENTARTCGSLCLN 1359
QY 888 GRIACSKVWCQPRE---CLLHKHSEC--PSGQSCPTILDDQCFVHPCTGVGBSCRSSL 941
Db 1360 GGTCTIS-----GPRSPTCLCLGSPFGPECPFPASSPCVG-----SNPCYNQGTCEPTSE 1408
QY 942 QP 943
Db 1409 SP 1410
RESULT 9
A40043
notch protein homolog TAN-1 precursor - human
C:Species: Homo sapiens (man)
C>Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 05-Oct-2004
C:Accession: A40043
R:Ellisen, L.W.; Bird, J.; West, D.C.; Soreng, A.L.; Reynolds, T.C.; Smith, S.D.; Sklar, Cell 66, 649-661, 1991
A:Title: TAN-1, the human homolog of the Drosophila Notch gene, is broken by chromosomal
A:Reference number: A40043; MUID:91347367; PMID:1831692
A:Accession: A40043
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: mRNA
A:Residues: 1-2555 <ELL>
A:Cross-references: UNIPARC:UPI0000177455; GB:M73980
C:Superfamily: notch protein; ankyrin repeat homolog; EGF homolog
F:261-292/Domain: EGF homolog <EGX1>
F:494-525/Domain: EGF homolog <EGF1>
F:987-1018/Domain: EGF homolog <EGK2>
F:1149-1180/Domain: EGF homolog <EGF3>
F:1187-1218/Domain: EGF homolog <EGF3>
F:1233-1264/Domain: EGF homolog <EGX3>
F:1927-1959/Domain: ankyrin repeat homolog <AN1>

F;1960-1992/Domain: ankyrin repeat homolog <AN2>
F;1994-2026/Domain: ankyrin repeat homolog <AN3>
F;2027-2059/Domain: ankyrin repeat homolog <AN4>
F;2060-2092/Domain: ankyrin repeat homolog <AN5>
Query Match 24.5%; Score 1532; DB 2; Length 2555;
Best Local Similarity 27.3%; Pred. No. 1.8e-80;
Matches 362; Conservative 125; Mismatches 365; Indels 472; Gaps 43;
QY 77 VCLKEYSQSVRTAGGPCSGSGSTPVIIGNTFNLKASGRNDRIRVLPPSFAMPRSRVTLV 136
Db 9 LCLALLPALAARGPRCS-QPGETCLNGG-----KCEAANGTEACVCGGAFVGRCO--- 58
QY 137 EAWDSNDTVPDSIIIEKASHSGMINPNSRQWOTLKONTGVVAHFEYQIRVTCDDYYVFGC 196
Db 59 -----DPNCLSTPCKNAGTCHVVDRRGV----- 82
QY 197 NKFCRPRDDFFGHYACQDQNGNKTCEGWMGPEC-----NRAI---CROGSPKHGSKCL- 247
Db 83 -----ADYAC-----SCALGFSGLCLTPLDNACLTNPCRNG-----GTCDLL 120
QY 248 ---PGRCROYGHQGLYC---DKCI PHPGCVHGI CNEPWO---CLCETNAGGQLCDKDL 297
Db 121 TLTEYKRCRCPPGWSGSKSCQOADPCASNP-CANGGQCLPFEASYICHCPSPFHGPTCRQDV 179
QY 298 NYCGTHQP--CLNGGTCSTNTPDKYOCSPGEGYSGNCEIAEHACLSDPCNHRGSKCT- 354
Db 180 NECG-QKPLRCRGGTCHNE-VGSRVCVCRATHGNCERPPYVPCSPQNGGTCRPTG 237
QY 355 SLGFCECECPGWTGPTCTSTNIDPCSPNNCSHGTCQDLVNGFKVCPCPQWTKTQOLDAN 414
Db 238 DVTHECACLPFGTGQNCENIDDCPGNCKNGGACVDGVNTYNCPCPPEWTGYCTEDVD 297
QY 415 ECE----- 417
Db 298 ECOLMPNACQNGTCKNTHGGYNCVVCVNGWTGBDCSENIDDCASACFHGATCHDRVASF 357
QY 418 ----- 417
Db 358 YCSCPGRGTLLCHLNDACISNPCNBSGNDTNPVNGKALCTCPGYTGTPACSDVDDECS 417
QY 418 --AKPCVNAKSKNLASYYYCDCLPGWMQNCIDINCLGQ--CONDASCRDLVNGYRCI 474
Db 418 LGANPEHAGKCNLTLSFECQCLQGYTGPRCEIDVNECVSNPCQNDATCLDQIGEPQCM 477
QY 475 CPPGYAGDHCDERDIDECASNPCLNGHHCQNEINRFPCLCPTGFGSNLQCLDIDYCEPNPC 534
Db 478 CMPGYGVHCEVNTDECASSPCLHNGRCLDKINEFOCECPTGTGTHLCO-DVDECASTPC 536
QY 535 QNGAQCVNRASDYFKCPEDYEGKNCSHLKHCRTPPCEVIDSCTVAMAS----- 584
Db 537 KNGAKCLDGPNTYTCVCTEGYTGTHCEVDIDECDDPDPCH-YGSKDGVATFTCLCRPGYT 595
QY 585 -NDTPGEVRYISSNVV-----GPH-----GKCK 606
Db 596 GHCEITNINESSQPCRELRTGTCQDPDNYALCFCLKGTTFGNCEINLDDCASSPCDSGTCL 655
QY 607 SQSGKFTCDNKGFTQTYCHEINDCESNPNCRNGGTCIDGV----- 648
Db 656 DKIDG-YECACERGYTGSNCSNIDECAGNPNCHNGGTCEDGNGFTCRCEPGYHDPCTCLS 714
QY 649 -----NSYKICSDGWEGAYCETNINDSQNPCHNGGTCDRLVNDV 689
Db 715 EVNECNSNPCHVAGCRDNLNGYKCDPDPGWSGTNCDINNNECESNPNVNGTCKDWTSGI 774
QY 690 YCDCKNGWKGTCHSRDSQCDDEATCNGGTCYDEGDAFKCMCPGMEGTTCTIARNSCL 749
Db 775 VCTCREGFSGNPCQTNINECASNPCLNKGTICIDDVAGYKCNCLLPTYGATCEVVL-APCA 833
QY 750 PNPCHNGGTC--VNGSGFTCV----- 769
Db 834 PSPCRNGGCRQSEDYESFCVCTAGAKGTCEVDINECVLSPCRHGASCONTHGYRC 893

770 -CKEWEPTCAQNTNDCSPHPCVNSGTCVGDWNYCEBAPGAPGDPDRINECQSSP 828
 894 HQQAGYSGRNCETDIDDREPNCHNGGCTDGINTAFCDCPLPGFRGTFCEDINECASDP 953
 829 CAFGA-----TCVDEINGYRCVCPGGH 850
 954 CRNGANCTCDVSDYTCCTCPAGFSGIHCENNTPTDCTESSCFNGGTCVGDGINSFCLCPGGF 1013
 851 SGAKCQ-----EVSGRPCTITWGSVIPDG-----AKWDDDCNTCQCL 886
 1014 TGSYQHVNECDSPCL-LGGTCQDGRGLHRCCTCPQGYTGPNQNLVHW---CDSSPCK 1069
 887 NG-----RIACKVWC-----PRCLLHKG----- 907
 1070 NGGCKWQTHQYRCFSPGWTGLYCDVPSVCEVAAQROGVQVVARLC-QHGGLCVDAGNT 1128
 908 -HSECPGQ--SCPIILDQCFVHPCGTVGCESSSLQPVKTKTSDSYVDNCANITFT 964
 1129 HHCRCQAGYTSVCEDLVSECSFPCQNGATC-TDYLGGYCKVA-GYHGVNCSBEI-- 1184
 965 FNKEMSPGVTTEHCSELRNILKNVSAEYSIYIACEPSPSANNIEHVAISAEDIRD 1024
 1185 --DECLSHPCQNGTCLDLEN-----TVKSC--PRGTQGVHCEINVDNCP 1228
 1025 GNPI 1028
 1229 VDPV 1232

RESULT 10
 S78549
 notch3 protein - human
 C;Species: Homo sapiens (man)
 C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
 C;Accession: S78549; S71825
 R;Joutel, A.; Tournier-Lasserre, E.
 submitted to the EMBL Data Library, April 1997
 A;Reference number: S78549
 A;Accession: S78549
 A;Molecule type: mRNA
 A;Residues: 1-2321 <JOU1>
 A;Cross-references: UNIPROT:Q9UM47; UNIPARC:UPI000011D827; EMBL:U97669; NID:q2668591; P
 R;Joutel, A.; Corpechot, C.; Ducros, A.; Vahedi, K.; Chabriat, H.; Mouton, P.; Alamowit
 x, M.M.; Weissenbach, J.; Bach, J.F.; Bousser, M.G.; Tournier-Lasserre, E.
 Nature 383, 707-710, 1996
 A;Title: Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke
 A;Reference number: S71825; MUID:97032728; PMID:8878478
 A;Accession: S71825
 A;Status: nucleic acid sequence not shown
 A;Molecule type: DNA
 A;Residues: 67-113;138-194;268-333,'G',335-346;536-613;716-765;1240-1279;1815-1888 <JOU2
 A;Cross-references: UNIPARC:UPI0000177457; UNIPARC:UPI0000177458; UNIPARC:UPI0000177459;
 C;Genetics:
 A;Gene: notch3
 A;Map position: 19p13.1
 C;Function:
 A;Description: may be involved in pathogenesis of CADASIL, causing a type of stroke and
 C;Superfamily: notch protein; ankyrin repeat homology; EGF homology
 C;Keywords: tandem repeat; transmembrane protein
 P;123-155/Domain: EGF homology <EGX1>
 F;162-194/Domain: EGF homology <EGF1>
 F;240-271/Domain: EGF homology <EGX2>
 F;318-349/Domain: EGF homology <EGX2>
 F;473-504/Domain: EGF homology <EGX3>
 F;853-884/Domain: EGF homology <EGF3>
 P;928-959/Domain: EGF homology <EGF4>
 F;1838-1870/Domain: ankyrin repeat homology <AN1>
 F;1871-1903/Domain: ankyrin repeat homology <AN2>
 F;1905-1937/Domain: ankyrin repeat homology <AN3>
 F;1938-1970/Domain: ankyrin repeat homology <AN4>
 P;1971-2003/Domain: ankyrin repeat homology <AN5>

Query Match 24.0%; Score 1498.5; DB 2; Length 2321;

Best Local Similarity 28.9%; Pred. No. 1.4e-78;
 Matches 324; Conservative 114; Mismatches 328; Indels 357; Gaps 37;
 QY 220 CMEGWNGPECN-RAICRQCSPPKHGCK---LPG-----DCRCQYQWGLYC---DKCIPH 268
 DB 67 CPGWNGERCOLEDPCHSFGPCAGRCVQSSVAGTARFSCRCPRGFRGPDCLSLPCLSS 126
 QY 269 PGCVHGI-C-----NEPWOCLCETNWGQQLCDKLNLYCGTHQPCNLNGTGTCSNTGPKYQCS 323
 DB 127 P-CAHGARSVGDGRFLCSCPPGYQGRSCRSVDCECRVGEPRHGGTCLNT-PGSFRQC 184
 QY 324 CPBGYSBPCEIAEHAELCLSDPCNRSCKET-SLGFCECSRGWGTCTCTNDDCSPNN 382
 DB 185 CPAGYTGPLENPAVPAPCAPSPCRNGGTCROSGDLTYDCACLPGFEGONCEVNVDDCPGHR 244
 QY 383 CSRGGTCQDLVNGFKVCPQWTGKTQCLDANCEAKP--CVNAKSCKNLIASYCQCLP 440
 DB 245 CLNGGTCVGVNYYNQCPPEWTGQCTEDVDECOQPNACHNGGTCTNLTGGHSCVGVN 304
 QY 441 GMMQNCNDINIDL----- 455
 DB 305 GMTGBSCSNIDDCATAVCFHGATCHDRVASFYCACPMGKTGLLCHLDDACVSNPCHEDA 364
 QY 456 ----- 455
 DB 365 ICTNPVNGRAICTCPGFTGGACDDQVDECSGANPCEHLGRVCNTQGSFLOCGRGYT 424
 QY 456 -----GQCNDASCRDLVNGYRCICPPGYAGDHCBDRDIDECASNPCLANGHC 502
 DB 425 GPRCETDVNECLSGPCRNQATCLDRIGQFTICWAGFTGYCEVDIDECOSSFCVNGVC 484
 QY 503 QNEINPQCLPTGFGSNLCQLDIDYCEBNPCQNGAQYNNRASYDFKCPDEDEYEGKNCVSH 562
 DB 485 KDRVNGFSCTCPGSGTSCQLDQVDECASTPCRNKAKCVQDPDGYECRCAEGEGTLCDR 544
 QY 563 LKDHCRTPC---EVID-----SCTVAMASNDI--PEGVYIISNVCVGHKCKSSGGK 612
 DB 545 NVDDCSPDPCHHGRGVDCDGIASFACAPGYGTGRCSQVDECRSQPRHGGKLDLV-DK 603
 QY 613 FTCDCKNGFTGTCHENINDCESNCRNGGTCIDGVNSYKICISDGEWAGYCTETNINDCS 672
 DB 604 YLCRCPSGTYGVNCEVNDICASNPC-TFVCDGDNIRYDCVQPGFTGLCNVEINECA 662
 QY 673 QNPCHNGGTCRDLVN----- 695
 DB 663 SSPCGEGGSCVDGNGFRCLCPPLPSHPHPCNHEPCSHGICYDAPGGFRVCCEP 722
 QY 696 GWKGTCH---SRDSQDEATCNGGTCYDDEGDAFKCMCPGGWGTCTNARNSSCLPNP 752
 DB 723 GMSGPRCSQSLARDA-CESQPCRAGTCSDDGMGFHCTCPPGVQGRQCELL--SPCTPNP 779
 QY 753 CHNGGTC-VVNGR----- 772
 DB 780 CEHGRCEAPGQLPVCSCEPQHQGRQCDQVDECAAPGAPGPHGICTNLGAFSCTCHG 839
 QY 773 GWEGPTCAQNTNDCSPHPCVNSGTCVGDWNYCECAAPGAPGDPDRININECOSSPCA-- 830
 DB 840 GYTPSCDQDINDCDPNCLNGGSCQDGVGSFSCSLPGFAGPCARDVDECLSNPCGPG 899
 QY 831 -----FGATCVDINGYRCVCPHSGAKC 855
 DB 900 TCTDHVASFTCTCPGYGFGHCEQDLPCDSPSSCFNGGTCVGVNSFSLCRPGYTGAHC 959
 QY 856 QEVS-----GRPCITMG--SVIPDGAkWDDCNTQCCLNGRIA--CSKV--WCSPRP--- 901
 DB 960 QHEADPCLSRPCLGGVCSAAPHGR-----CTLESFTGPGCQTLVDMCSRQPCQNG 1012
 QY 902 --CLLKHGSECP---SGQSC-----IPILD----- 922
 DB 1013 GRVCVQTGAYCLCPGPGSGRLCIRSLPCREAAAIQVRLQLQAGQGCQVDESSHYCVC 1072
 QY 923 -----DQCFVHPCTGVGECRSSLOPVTKTCTSDSYODNCANITFTFNKEM 969

Db 1073 PEGRSGHCEQVDPCLAPQCHQGTCTCR-GYMGYMCECL-PGVNGDNCEDDVDECAQSP 1130
Qy 970 MSPG-----LTTTHICS-EURNLNILKNVSAEYSIYIACEPSP 1006
Db 1131 CORHGGSCIDILVARYLCSPPGTGLVLCINEED-----DCGPGP 1168
RESULT 11
T31070
notch homolog - sea urchin (Lytechinus variegatus)
C;Species: Lytechinus variegatus (variegated urchin)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 31-Jan-2000
C;Accession: T31070
R;Sherwood, D.R.; McClay, D.R.
Development 124, 3363-3374, 1997
A;Title: Identification and localization of a sea urchin Notch homologue: insights into
A;Reference number: Z20966; MUID:97454256; PMID:9310331
A;Accession: T31070
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-2531 <SHE>
A;Cross-references: UNIPARC:UPI000007E31C; EMBL:AF000634; NID:g2570350; PID:g2570351; PI
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology
Query Match 23.9%; Score 1495.5; DB 2; Length 2531;
Best Local Similarity 27.8%; Pred. No. 2.2e-78; Indels 357; Gaps 36;
Matches 334; Conservative 136; Mismatches 376;
Qy 51 NGNCC-----GGARNPGDRKTRDECDTYFFVKLCIKEYOSRVTAGGPCSFGSGSTPVIGNT 106
Db 213 NGGTCFNQYGGFCCEPLGFTGQCELVYBPCSPD-----PCRNG-GQCASTGPT 262
Qy 107 PNLKASGNDRNRIVLFPFSFAMPRSYTLLEAVDSSNDTVQPSIIEKASHGMINPSRQ 166
Db 263 FTCTCQNGFTGETCEL-----NLNDCTQHCLNGGTCIDGVNDYT-- 302
Qy 167 WQTLKQNTGV-AHPEY-QIRVTCDDYYGFGCKNFCRPRDDFFGHYACDQNGNKTMEGW 224
Db 303 CSCLEKFTGYCEMDFDECDAVDPCFNGGTCS-----NTYGNFSC-----ICVRGW 349
Qy 225 MGPECNRAICRQGSCKPHGSKLPGDQR-----COYGWQGLYC---DKCIPHPGCV 272
Db 350 EGQTC E--INKDDCTP--NPQFGECEDRVASPKCTCPGRTGLLCHLEDACMSNP-CH 404
Qy 273 H-----GICNE-----PWQCLCE 285
Db 405 HTAQCTSVVDGSPICDCAATGQYGFNGSEDIDECSSLMSDSTICQSGGTQCNFPGGWSCLCS 464
Qy 286 TNNWGQLCDLNYCGTHQPCLANGTCSNTGPDKYQCSCEPGYSGPNCEIAEHAACLDPC 345
Db 465 SGFTGSRCETDIDEC-DDDPCYNGGTCLNK-RGGYACILTGFTGTJCTETDINECSSNPC 522
Qy 346 HNRGSKETSILGFCEGSPGWGTCTSTNIDDCSPNNCSHGTCQDLVNGFKVCYCPQWT 405
Db 523 LNGASCFDITGRFECACLAGYGTTCQVNIIDDCQSSPCENGTCIDGVNQFTCLCETGYE 582
Qy 406 GKTQCLDANECEAKPCVNAKSKNLIASYYCDLPGMWQNCNINIDCLGQ-QONPASC 464
Db 583 GHRCEMDSDECAFRPCNMGGVCEDLIGFYQCNCPVGTSGDNCEYNHYDCSSNPNVNDGTC 642
Qy 465 RDLVNGYRCICPPGVAGDHCEERDIDECASNPLANGHCONEINFCQICPTGFRSGNLQOL 524
Db 643 VDGINEYTCMEHGTGRLNCEEDIDDCESRPNCHNGGTCTVDVNGYHCLCPGYHDPFMS 702
Qy 525 DIDYCEPNPQNGAQCYNRASDYFKCPEDYEGKNCNHLKDHCRTPCE----- 573
Db 703 NINECSSNPNVNGGSGCHDVNGVEYCECMAGYGTGTRCTDDDFDECSSNPQCHGGTCDNRHAF 762
Qy 574 -----VIDSCTVAMSN-----DTPGVRVY-----S 595
Db 763 YNCTCOAGYTGINCENVIDDCVDFPLNGGTCIDEVNSFCQCPQTFFVLLCETERSPCE 822
Qy 596 SNVCGPHGKC-KSQSGGKFTCDCKNGFTGTGTCHENINDCESNPNCRNGGTCIDGVNSYKCI 654

Db 823 DNOQNGATCVYSEDYAGYSCRTSGFQGNFCDDDRNECLFSPCRNGSGCTNLEGSFCS 882
Qy 655 CSDGWEGAGVCEYTNINDCSNPCHNGTCTRDLVNDFYCDCKXGHWKTKTCHSRDQCEATC 714
Db 883 CLPGYDGPICEINIDRECAAGPCTNGGICTDLIDDFYCSCQGRFTGRKNQNDTDECLSSPC 942
Qy 715 NNGGTCYDSDGDAFKMCPGSGWEGTTCNI-----ARNSSCL----- 749
Db 943 RRGATCHEYVDYSYTCCLVGFSGMHCEINDQDCTTSCLYGGTCIDGVNSYTCCEVTGYT 1002
Qy 750 -----PNPCHNGG 757
Db 1003 GSNQCIETINECDSDPCENGATCQDRFGSYSCHDVGFTGLNCEHVYVQWCSQNNPCVNGA 1062
Qy 758 TCVVNG-----ESFTCV 769
Db 1063 TCVAMGHLYECHCASNNWIKGLDVPKVSCDIAASDKNVRSELCLNGGTCIDATSSHCL 1122
Qy 770 CKEGWEGPICAOQNTNDCSPHPCVNSGTCVDGDMWYRCECAPGFAGPDCRININECOSSPC 829
Db 1123 CQDGYTGSYCEVNIDECASAPCHNGGTCTDGVSYTCSCLPFGFEGPRCQONINECASSPC 1182
Qy 830 AFGATCVDIEINGYRCVCPPGHSGAKQGVSGRCPITMGSVIPDGAKWDDDCNTCQCLNGR 889
Db 1183 HNGGQCHDMVNGYTCSCPAQTGTDCS-----INL-----DDCYEGACYHGG 1224
Qy 890 IACSKVWCGRRPCLLKHGSECSGSCIPILDDQCFV-HPCTG-VCECHSSSLQPVKTK 947
Db 1225 VCIDQV--GTYTC-----DCPLG-----FVGQHCEDGVNECLSNPCDPVGSQ 1264
Qy 948 CTSDSYVDNCANITFTFNKEMWSPGLTTEHCISELNLNLKNVSAEYSIYIACEPSPS 1007
Db 1265 -----DCVQLINNY-OCVKRPGYTGQDCQEIPN-----CQNDPC 1298
Qy 1008 ANN 1010
Db 1299 QNN 1301
RESULT 12
T30201
Notch homolog protein - sea squirt (Halocynthia roretzi)
C;Species: Halocynthia roretzi
C;Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 08-Sep-2002
R;Hori, S.; Saitoh, T.; Matsumoto, M.; Makabe, K.W.; Nishida, H.
Dev. Genes Evol. 207, 371-380, 1997
A;Title: Notch homologue from Halocynthia roretzi is preferentially expressed in the cer
A;Reference number: Z20775
A;Accession: T30201
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-2352 <HOR>
A;Cross-references: UNIPARC:UPI0000081BFC; EMBL:AB001327; NID:d1204472; PID:d1026501; PI
C;Genetics:
A;Gene: Notch
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology
Query Match 23.9%; Score 1492; DB 2; Length 2352;
Best Local Similarity 28.7%; Pred. No. 3.3e-78;
Matches 370; Conservative 126; Mismatches 409; Indels 382; Gaps 45;
Qy 49 LONGNCCGGARNPGD--RKCTRDECDYF--KVCLEKYQSRVTAGGSCFSGSGSTPVIGG 104
Db 77 LNSGTCVAVTELTAAGDFSIQCT---CQTGTGDTCT---SQVLVYCSNPNCSNGAGCEEL--S 128
Qy 105 NTFNLKASRG-----NDRNRIVL-----FSPAWPRSYTLVVEAWDSS 142
Db 129 NSFKCTCTSGYGTCTCANDVNECDTPDICQVAGTCSNNDGGYSCS-----CVAGFEGN 181
Qy 143 NDTVPQDSIIIEKASHSG-----MINPSQWQTLK 171


```
Qy 671 CSQPNCHNGGTCRDLVNDYF-----CDC 693
  |||:|||||:
Db 662 CASSPCGEGSGVDGNGFHCCLCPGSLPPLCLPANHPCAHKPCSHGVCHDAPGGPRVC 721
  |||:|||||:
Qy 694 KNGHKGTCHS--RDSQCDKATCNNGTCVDEGDAFKCMCPGGWEGTTCNIARNSSCLPN 751
  |||:|||||:
Db 722 EPGSGPRCSQSLAPDACESQPCOAGGTCTSDGIGFRCTCAPGQHQCEVL--SPCTPS 779
  |||:|||||:
Qy 752 PCHNG-----GTCVVNGESFTCVCK 771
  |||:|||||:
Db 780 LCEHGHCESDPDLITVCSPPHGWQRPCQDVDEACAGSPCGPHGTCNLPGNFRCLCH 839
  |||:|||||:
Qy 772 EGMEGPTCAQNTNDCSPHPVCNSGTCTVDGDNWYRCECAPGAGPDCRININECOSSPCAF 831
  |||:|||||:
Db 840 RGYTGPPCDQIDDCDNPCLHGSCQDGVGSPFSCCLDGFAGPRCARDVDECLSSPCGP 899
  |||:|||||:
Qy 832 GATCVDEINGYRCVCPGHSGAKCQ---EVSGRPCITMGSVIPDGAKWD----- 877
  |||:|||||:
Db 900 G-TCTDHSVAFTCACPPGYGFGHCEIDLPCSPSPSCFNGGTCVDGVSSFSCLCRPGYGTG 958
  |||:|||||:
Qy 878 -----DDCNTCCOCLNGRIACSKV-----WCGPRPC----- 902
  |||:|||||:
Db 959 HCQYADPCFSRPLCHGI-CNPHPGFECTREGFTGSOQNPVDCSQAPCONGRCV 1017
  |||:|||||:
Qy 903 -----LLKHGHE---CPSGQ 915
  |||:|||||:
Db 1018 QTGAYCICPPGWSRLCDIQSLPCTEAAAQWVLEQLCQEGGKCIDKGRSHYCVCPCEGR 1077
  |||:|||||:
Qy 916 S---CIPILDQCFVHCTGVGBRSSSLQPVTKTSDTSYYQDNCA-NITFTFNKEMMS 971
  |||:|||||:
Db 1078 TGSHEHEV-DPTAQPCQHGTCR-GYMGYYVCECPA-GYAGDSCEBNDDECASQPCQN 1134
  |||:|||||:
Qy 972 PG---LTTBHCIS-ELRNILKNVSAEYSIVYIACPPSPSANNEL 1012
  |||:|||||:
Db 1135 GGSCLDLVARYLCSPPGTGLVLCNEIN-----DCDLGSLDSGV 1175
  |||:|||||:

RESULT 15
A49175
Motch B protein - mouse (fragment)
N;Alternate names: Motch homolog
C;Species: Mus musculus (house mouse)
C;Date: 21-Jan-1994 #sequence revision 05-Jan-1996 #text_change 05-Oct-2004
C;Accession: A49175; PH1570; S32113
R;Lardelli, M.; Lendahl, U.
Exp. Cell Res. 204, 364-372, 1993
A;Title: Motch A and Motch B--two mouse Notch homologues coexpressed in a wide variety of
A;Reference number: A49175; MUID:93178563; PMID:8440332
A;Accession: A49175
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-1203 <LAR>
A;Cross-references: UNIPROT:O35516; UNIPARC:UPT0000016CF1P; EMBL:X68279; NID:G287989; PII
A;Experimental source: embryo
A;Note: sequence extracted from NCBI backbone (NCBIP:126158)
C;Comment: This protein has many EGF repeats and lin-12/Notch repeats.
C;Comment: This protein is one of the neurogenic proteins controlling the decision betwe
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology
F;143-174/Domain: EGF homology <EGX1>
F;482-513/Domain: EGF homology <EGF1>
F;560-591/Domain: EGF homology <EGF>
F;674-705/Domain: EGF homology <EGX2>
F;712-743/Domain: EGF homology <EGF3>
F;836-867/Domain: EGF homology <EGX3>

Query Match 23.5%; Score 1467.5; DB 2; Length 1203;
Best Local Similarity 32.7%; Pred. No. 4.9e-77;
Matches 295; Conservative 113; Mismatches 316; Indels 179; Gaps 28;

Qy 208 GHYACDONGNKTCEHGMWGECHRAICRQCSFKHSGSKLPG-----DCRCQYGW 257
  |||:|||||:
Db 3 GGYGC-----VCVNGSGDSCSENI--DDCA--YASC-TPGSTCIDRVASFSLCPEKG 51
  |||:|||||:
```

Search completed: February 4, 2006, 09:36:35
Job time : 64 secs

```
Qy 258 QGLYC---DKCIHPHPCVHGIC-----NEPQCLCETNWGGQLCDKOLNYC--GTHOPCL 307
  |||:|||||:
Db 52 AGLLCHLDDACISNPCHKGALCDTNPLNGQYICTCPQGYKGADCTEDVDCEAMANSNPE 111
  |||:|||||:
Qy 308 NGGTCSTNTPDKYQCSPEGYSGPNCEIAHACLSDPCHNRGSGKETSILGPECSCSPWT 367
  |||:|||||:
Db 112 HAGKCVNT-DGAFHCEBLKGYAGPRCEBMDINECHSDPCQNDATCLDGIKGGFTCLCMGPK 170
  |||:|||||:
Qy 368 GPTCSTNIDDCSPNNSHGCTCODLVNGFKVCPPOWTKTQCLDANECBAKPCVNAKSC 427
  |||:|||||:
Db 171 GVHCELEVNQCQSNPCVNNQCVDKVNRFQCLCPGFTGPVCQIDIDCSSTPCLNGAKC 230
  |||:|||||:
Qy 428 -----KNL-----IASYCDCLPQWMMQNCIN 450
  |||:|||||:
Db 231 IDHPNGYECQCATGFTGILCDENIDNCDDPCHHGQCODGIDSYTCICNPGYMGAI CSDQ 290
  |||:|||||:
Qy 451 INCLGQ-CONDASCDRLVNGYRCICPPGYAGDHCEBIDEACSNPCLNGHGHQNEINRF 509
  |||:|||||:
Db 291 IDECYSSPCLNDGRCIDLVNGYQCNCQPGTSGNLNFDDCASNPCMH-GVCVDGINRY 349
  |||:|||||:
Qy 510 QCLCPTGFSNLCOLDIDYCEPNPCONGAOCYNRASDYFCCKPEDYEGKNCSHLKHCR 569
  |||:|||||:
Db 350 SCVSPGFTGRCNIDIDEASNPCKGATCINDVNGFRICIEGPHHPSCYSQVNECLS 409
  |||:|||||:
Qy 570 TPCEVIDSCTVAMASNDTPGVRYIS-----SNVCGPHGKCKSQSGKFTCDN 618
  |||:|||||:
Db 410 NPC-IHGNCTGGLSGYKCLCDAGWGVNCEVDKNECLSNPCNGGT CNLVNG-YRCTCK 467
  |||:|||||:
Qy 619 KGFTGYTCHENINDCSNPNCRNGGTCIDGVN----- 649
  |||:|||||:
Db 468 KGFKYGNQVNIIDECASNPCLNQGTCTFDDVSGYTCMCLPYTGKNCQTVLAPCSNPEN 527
  |||:|||||:
Qy 650 -----SYKICSDGMEGAYCETNIDCSQNPCHNGGTCRDLVNDPYCDCKNGWK 698
  |||:|||||:
Db 528 AAVCKEAPNPFESCLCAPGWQGRCTVDDECLSKPCMNNGVCHNTQGSYVCECPGFS 587
  |||:|||||:
Qy 699 GKTCHSRDSQCBATCNNGGTCYDEGDAFKCMCPGGWEGTTCNIARNSSCLPNPCHNGGT 758
  |||:|||||:
Db 588 GMDCEBDINDCLANPCQNGSGSCVDHVNFTFSCQCHPGFIGDKCQTDMM-ECLSEFCNKG 646
  |||:|||||:
Qy 759 CVVNGESFTCVCKEGWEGPICAQNTNDCSPHPVCNSGTCTVDGDNWYRCECAPGAGPDCR 818
  |||:|||||:
Db 647 CSDYVANSYTCCTCPAGFHGVHCENNIDECTESSCFNGGTCVDGINSFSCLCPPVGTGPFCL 706
  |||:|||||:
Qy 819 ININECOSSPCAFGATCVDSEINGYRCVCPGPHSGAKCOEV-----SGRPCITMGSVIPDGA 874
  |||:|||||:
Db 707 HDINECSSNPLNAGTCVDGLGYTRCICPIGYTKNCQTLVNLCSRSPCKNKGTCVQDKA 766
  |||:|||||:
Qy 875 K-----WDDDCNTCCQLNGRIACSKVWCGPRPCLLHKG----- 907
  |||:|||||:
Db 767 RPHCLCPPGWDG--AYCDVLN--VSC-----KAAALQKGVPEVHLQHSIGICINAGNT 815
  |||:|||||:
Qy 908 -HSECP---SQGSCIPILDQCFVHCTGVGBRSSSLQPVTKC-----TSDSYQDN 957
  |||:|||||:
Db 816 HHQCQPLGYTGSYCEBQL-DEACSNPCQHGATC-NDFIGGYRCCEVPGYGVNCEYEVD 873
  |||:|||||:
Qy 958 CAN 960
  |||:|||||:
Db 874 CQN 876
  |||:|||||:
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 09:35:44 ; Search time 52 Seconds
(without alignments)
1696.441 Million cell updates/sec

Title: US-10-650-650-18

Perfect score: 6248

Sequence: 1 MRSPTGRSGRPLSLLLAL.....LIAVAEVRQRRPLKNRTD 1067

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : IssuedPatentSeqs*

- 1: /cgn2_6/ptodata/1/iaa/5_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/6_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/H_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6248	100.0	1067	2	US-09-579-536C-18
2	6248	100.0	1218	2	US-08-882-046-2
3	6248	100.0	1218	2	US-09-068-740A-11
4	6248	100.0	1218	2	US-09-566-047-2
5	6248	100.0	1218	2	US-09-579-536C-1
6	6248	100.0	1218	2	US-09-949-016-5902
7	6248	100.0	1254	2	US-09-949-016-10297
8	6241	99.9	1218	2	US-09-917-254-85
9	6233	99.8	1218	1	US-08-400-159-6
10	6233	99.8	1218	2	US-08-611-729A-6
11	6233	99.8	1218	2	US-09-195-524-6
12	6233	99.8	1218	2	US-09-310-685-4
13	6228	99.7	1218	2	US-09-214-278-7
14	6228	99.7	1218	2	US-09-855-722-7
15	6176	98.8	1208	2	US-09-199-865-1
16	6176	98.8	1208	2	US-10-213-329-1
17	6093	97.5	1036	2	US-09-068-740A-6
18	6093	97.5	1187	2	US-09-068-740A-7
19	6089	97.5	1219	2	US-08-882-046-5
20	6089	97.5	1219	2	US-09-566-047-5
21	5969	95.5	1010	2	US-08-882-046-7
22	5969	95.5	1010	2	US-09-566-047-7
23	5466.5	87.5	1193	1	US-08-400-159-10
24	5466.5	87.5	1193	2	US-08-611-729A-10
25	5466.5	87.5	1193	2	US-09-195-524-10
26	5466.5	87.5	1193	2	US-09-310-685-8
27	3675	58.8	1238	2	US-09-214-278-5

28	3675	58.8	1238	2	US-09-855-722-5	Sequence 5, Appli
29	3662	58.6	1212	2	US-09-214-278-3	Sequence 3, Appli
30	3662	58.6	1212	2	US-09-855-722-3	Sequence 3, Appli
31	3654	58.5	1055	2	US-09-214-278-2	Sequence 2, Appli
32	3654	58.5	1055	2	US-09-855-722-2	Sequence 2, Appli
33	3631.5	58.1	1248	2	US-08-882-046-6	Sequence 6, Appli
34	3631.5	58.1	1248	2	US-09-566-047-6	Sequence 6, Appli
35	3346	53.6	1148	2	US-08-882-046-4	Sequence 4, Appli
36	3346	53.6	1148	2	US-09-566-047-4	Sequence 4, Appli
37	3291	52.7	1065	1	US-08-400-159-8	Sequence 8, Appli
38	3287	52.6	1257	2	US-08-611-729A-8	Sequence 8, Appli
39	3287	52.6	1257	2	US-09-195-524-8	Sequence 8, Appli
40	3287	52.6	1257	2	US-09-310-685-6	Sequence 6, Appli
41	2339.5	37.4	1404	1	US-08-400-159-2	Sequence 2, Appli
42	2339.5	37.4	1404	2	US-08-611-729A-2	Sequence 2, Appli
43	2339.5	37.4	1404	2	US-09-195-524-2	Sequence 2, Appli
44	1585.5	25.4	2523	1	US-08-185-432-18	Sequence 18, Appli
45	1585.5	25.4	2523	2	US-08-899-232-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-579-536C-18

; Sequence 18, Application US/09579536C

; Patent No. 6716974

; GENERAL INFORMATION:

; APPLICANT: MACIAG, Thomas

; APPLICANT: ZIMRIN, Ann

; APPLICANT: SWALLA, Deena

; APPLICANT: PRUDOVSKI, Igor

; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC METHODS AND COMPOSITIONS BASED ON JAGG

; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS

; FILE REFERENCE: 053689-5002-01

; CURRENT APPLICATION NUMBER: US/09/579,536C

; CURRENT FILING DATE: 2000-05-24

; PRIOR APPLICATION NUMBER: US 09/199,865

; PRIOR FILING DATE: 1998-11-25

; PRIOR APPLICATION NUMBER: PCT/US97/09407

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/018,841

; PRIOR FILING DATE: 1996-05-31

; NUMBER OF SEQ ID NOS: 56

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 18

; LENGTH: 1067

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-579-536C-18

Query Match 100.0%; Score 6248; DB 2; Length 1067;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1067; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MRSPTGRSGRPLSLLLALCALRAKVCAGSOFLEILSMQNVNGLQNGCCGARN	60
Db	1	MRSPTGRSGRPLSLLLALCALRAKVCAGSOFLEILSMQNVNGLQNGCCGARN	60
Qy	61	PGDRKCTRDECDYFKVCLKEYOSRVTAGGPGSGSTVIGGNTFNKASRGNDNRRI	120
Db	61	PGDRKCTRDECDYFKVCLKEYOSRVTAGGPGSGSTVIGGNTFNKASRGNDNRRI	120
Qy	121	VLPFSFAWPRSYTLLVEAMDSSNDTVQPSIIIEKASHSGMINPSROWOTLKQNTGVAHFE	180
Db	121	VLPFSFAWPRSYTLLVEAMDSSNDTVQPSIIIEKASHSGMINPSROWOTLKQNTGVAHFE	180
Qy	181	YQIRVTCDYYGFGCNKFCRPRDDFFGHYACDQNGKTCMEGMWGPCNRAICRQCSP	240
Db	181	YQIRVTCDYYGFGCNKFCRPRDDFFGHYACDQNGKTCMEGMWGPCNRAICRQCSP	240
Qy	241	KHGSCKLPDGCRCQYWGQGLYCDKCTPHPCVHGICNEPWQCICETWNGQOLCDKDLNYC	300
Db	241	KHGSCKLPDGCRCQYWGQGLYCDKCTPHPCVHGICNEPWQCICETWNGQOLCDKDLNYC	300

241 KHGSKLPGDCRCOYQWGLYCDKCIHPGCVHGINCPWQCLCETNMGQCLDKDLYNC 300
301 GTHQPCNLNGTCSNTGPDKYQCSPEGYSGPNCIEAIAHACLSDPCHNRGSKETSLSGFEC 360
301 GTHQPCNLNGTCSNTGPDKYQCSPEGYSGPNCIEAIAHACLSDPCHNRGSKETSLSGFEC 360
361 ECSPGWTGPTCSNIDDCSPNCSHGCTCDLVNGFKVCVCPQWTKTCDLDANECEAKP 420
361 ECSPGWTGPTCSNIDDCSPNCSHGCTCDLVNGFKVCVCPQWTKTCDLDANECEAKP 420
421 CVNAKSKNLIASYYCDCLPGWQNGQCDININDCLGQCDNDASCRDLVNGYRCICPPGYA 480
421 CVNAKSKNLIASYYCDCLPGWQNGQCDININDCLGQCDNDASCRDLVNGYRCICPPGYA 480
481 GDHCERDIDECASNPLNGHGHQNEINRFQCLPTGFGSGLCOLDIDYCEPNPQNGAQC 540
481 GDHCERDIDECASNPLNGHGHQNEINRFQCLPTGFGSGLCOLDIDYCEPNPQNGAQC 540
541 YNRASDYFCKPEDYEGKNCNCSHLKHCRTPCEVIDSCTVMAASNDTPEGVRYISSNVCG 600
541 YNRASDYFCKPEDYEGKNCNCSHLKHCRTPCEVIDSCTVMAASNDTPEGVRYISSNVCG 600
601 PHGKCKSQSGGKFTCDCKNGKFTGTGYCHENINDCESNPCRNGGTICDGVNSYKICSDGWE 660
601 PHGKCKSQSGGKFTCDCKNGKFTGTGYCHENINDCESNPCRNGGTICDGVNSYKICSDGWE 660
661 GAYCETNINDCSQNPCHNGGTCRDLVNDYFCDCKNGWKGTCHSRDSQCDDEATCNGGTC 720
661 GAYCETNINDCSQNPCHNGGTCRDLVNDYFCDCKNGWKGTCHSRDSQCDDEATCNGGTC 720
721 YDEGDAFKMCPGMEGTGTCNIAKNSCLNPNCHNGGTCVVGESFTVCCKEGWGPICA 780
721 YDEGDAFKMCPGMEGTGTCNIAKNSCLNPNCHNGGTCVVGESFTVCCKEGWGPICA 780
781 QNTNDCSPHPCYNSTGTCVGDGNWYRCAPAGPDCRININEQSSPCAFAGATCVDEIN 840
781 QNTNDCSPHPCYNSTGTCVGDGNWYRCAPAGPDCRININEQSSPCAFAGATCVDEIN 840
841 GYRCVCPGHSQACQEVSGRPTITMGSVIPDGAHAKDDDCNTCQCLNGRIACSKVWCGR 900
841 GYRCVCPGHSQACQEVSGRPTITMGSVIPDGAHAKDDDCNTCQCLNGRIACSKVWCGR 900
901 PCLLKHGSHCPGSCQICPILDDCFVHPTGVCGERSSSIQPVKTKTSDSYQDNCAN 960
901 PCLLKHGSHCPGSCQICPILDDCFVHPTGVCGERSSSIQPVKTKTSDSYQDNCAN 960
961 ITFTFNKEMSPGLTTEHICSELNMLKNVSAEYSIYIACEPSPSANNIEHVAISAED 1020
961 ITFTFNKEMSPGLTTEHICSELNMLKNVSAEYSIYIACEPSPSANNIEHVAISAED 1020
1021 IRDGNPIKEITDKIIDLVSRRDGNSSIIAAVAEVRVORRPLKNRTD 1067
1021 IRDGNPIKEITDKIIDLVSRRDGNSSIIAAVAEVRVORRPLKNRTD 1067

RESULT 2
US-08-882-046-2
Sequence 2, Application US/08882046
Patent No. 6136952
GENERAL INFORMATION:
APPLICANT: Li, Linheng
APPLICANT: Hood, Leroy
APPLICANT: Krantz, Ian D.
APPLICANT: Spinner, Nancy B.
TITLE OF INVENTION: Human Jagged Polypeptide, Encoding
TITLE OF INVENTION: Nucleic Acids and Methods of Use
NUMBER OF SEQUENCES: 110
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA

ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/882,046
FILING DATE: 25-JUN-1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-UW 2637
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1218 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-882-046-2
Query Match 100.0%; Score 6248; DB 2; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1067; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRSRPRTRGRSRLPLALLCALRAKVCAGSQFOLEILSMQNVNGLQNGCCGARN 60
DB 1 MRSRPRTRGRSRLPLALLCALRAKVCAGSQFOLEILSMQNVNGLQNGCCGARN 60
QY 61 PGRKCTRDECDTYFKVCLKEYQSRVTAGPCFSGSGSTPVI GGNTPNLKASGRNDRNI 120
DB 61 PGRKCTRDECDTYFKVCLKEYQSRVTAGPCFSGSGSTPVI GGNTPNLKASGRNDRNI 120
QY 121 VLPFSFAWPSRYTLLEAVNDSSNDTVQPDSEIKASHGMINPSROWOTLKONTGVAHFE 180
DB 121 VLPFSFAWPSRYTLLEAVNDSSNDTVQPDSEIKASHGMINPSROWOTLKONTGVAHFE 180
QY 181 YQIRVTCDYFYGGNCKFCRPRDDFFGHYACDQNGNKTCEMGWMPGECNRAICRQGCSP 240
DB 181 YQIRVTCDYFYGGNCKFCRPRDDFFGHYACDQNGNKTCEMGWMPGECNRAICRQGCSP 240
QY 241 KHGSKLPGDCRCOYQWGLYCDKCIHPGCVHGINCPWQCLCETNMGQCLDKDLYNC 300
DB 241 KHGSKLPGDCRCOYQWGLYCDKCIHPGCVHGINCPWQCLCETNMGQCLDKDLYNC 300
QY 301 GTHQPCNLNGTCSNTGPDKYQCSPEGYSGPNCIEAIAHACLSDPCHNRGSKETSLSGFEC 360
DB 301 GTHQPCNLNGTCSNTGPDKYQCSPEGYSGPNCIEAIAHACLSDPCHNRGSKETSLSGFEC 360
QY 361 ECSPGWTGPTCSNIDDCSPNCSHGCTCDLVNGFKVCVCPQWTKTCDLDANECEAKP 420
DB 361 ECSPGWTGPTCSNIDDCSPNCSHGCTCDLVNGFKVCVCPQWTKTCDLDANECEAKP 420
QY 421 CVNAKSKNLIASYYCDCLPGWQNGQCDININDCLGQCDNDASCRDLVNGYRCICPPGYA 480
DB 421 CVNAKSKNLIASYYCDCLPGWQNGQCDININDCLGQCDNDASCRDLVNGYRCICPPGYA 480
QY 481 GDHCERDIDECASNPLNGHGHQNEINRFQCLPTGFGSGLCOLDIDYCEPNPQNGAQC 540
DB 481 GDHCERDIDECASNPLNGHGHQNEINRFQCLPTGFGSGLCOLDIDYCEPNPQNGAQC 540
QY 541 YNRASDYFCKPEDYEGKNCNCSHLKHCRTPCEVIDSCTVMAASNDTPEGVRYISSNVCG 600
DB 541 YNRASDYFCKPEDYEGKNCNCSHLKHCRTPCEVIDSCTVMAASNDTPEGVRYISSNVCG 600
QY 601 PHGKCKSQSGGKFTCDCKNGKFTGTGYCHENINDCESNPCRNGGTICDGVNSYKICSDGWE 660
DB 601 PHGKCKSQSGGKFTCDCKNGKFTGTGYCHENINDCESNPCRNGGTICDGVNSYKICSDGWE 660

661 GAYCETNINDCSQNPCHNGGTCRDLVNDYFCDCKNGKWKTKCHSRDSCQDEATCNGGTC 720
661 GAYCETNINDCSQNPCHNGGTCRDLVNDYFCDCKNGKWKTKCHSRDSCQDEATCNGGTC 720
721 YDEGDAFKMCPGCGWEGTTCNIARNSSCLPNPCHNGGTCVVGSGESFTVCCKEGWEGPICA 780
721 YDEGDAFKMCPGCGWEGTTCNIARNSSCLPNPCHNGGTCVVGSGESFTVCCKEGWEGPICA 780
781 QNTNDCSHPHCYNSTGTCVGDGNWYRCECAPGAGPDCRININECQSSPCAFGATCVDDEIN 840
781 QNTNDCSHPHCYNSTGTCVGDGNWYRCECAPGAGPDCRININECQSSPCAFGATCVDDEIN 840
841 GYRCVCPGSHGAKCQVSGRPCTITMGSVIPDGAKWDDDCNTCOCLNGRIACSKVWCGR 900
841 GYRCVCPGSHGAKCQVSGRPCTITMGSVIPDGAKWDDDCNTCOCLNGRIACSKVWCGR 900
901 PCLLHKHSHSCPSQSCIPILDDQCFVHPCTGVGECRSSLSLPVKTKTSDSYQDNCA 960
901 PCLLHKHSHSCPSQSCIPILDDQCFVHPCTGVGECRSSLSLPVKTKTSDSYQDNCA 960
961 ITFTFNKEMMSPGLTTEHICSELRLNMLKNVSAEYSIYIACEPSPSANNHVAISAED 1020
961 ITFTFNKEMMSPGLTTEHICSELRLNMLKNVSAEYSIYIACEPSPSANNHVAISAED 1020
1021 IRDGNPIKEITDKIIDLVSKRDNSSSLIAAFAEVRVQRPLKNRTD 1067
1021 IRDGNPIKEITDKIIDLVSKRDNSSSLIAAFAEVRVQRPLKNRTD 1067

RESULT 3

US-09-068-740A-11
; Sequence 11, Application US/09068740A
; Patent No. 637387
; GENERAL INFORMATION:
; APPLICANT: SAKANO, SEIJI
; APPLICANT: ITOH, AKIRA
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8447
; CURRENT APPLICATION NUMBER: US/09/068,740A
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: JP 7-299611
; PRIOR FILING DATE: 1995-11-17
; PRIOR APPLICATION NUMBER: JP 7-311811
; PRIOR FILING DATE: 1995-11-30
; PRIOR APPLICATION NUMBER: PCT/JP96/03356
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 11
; LENGTH: 1218
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-068-740A-11

Query Match 100.0%; Score 6248; DB 2; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1067; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MRSPTRGSRPLSLILALCALRAKVCASQGFLEILSMQNVGELONGNCCGARN 60
1 MRSPTRGSRPLSLILALCALRAKVCASQGFLEILSMQNVGELONGNCCGARN 60
61 PGRDKTRDSCDYFYKCLKEYQSRVTAGPCFSGSGSTPVGNTFNLKASRGNDRNI 120
61 PGRDKTRDSCDYFYKCLKEYQSRVTAGPCFSGSGSTPVGNTFNLKASRGNDRNI 120
121 VLPFSFAPWRSYTLLEAWDSNDTVPDSDIIKASHSGMINPSRQWTLKQNTGVAFHE 180
121 VLPFSFAPWRSYTLLEAWDSNDTVPDSDIIKASHSGMINPSRQWTLKQNTGVAFHE 180
181 YQIRVTCDDYYFGCNKFCRPRDDPFHGVACDQNGKTCMEGWMGPECNRAICROGCSP 240
181 YQIRVTCDDYYFGCNKFCRPRDDPFHGVACDQNGKTCMEGWMGPECNRAICROGCSP 240

241 KHGSKCLPGDCRCQYQGQGLYCDKCIPIHPGCVHGINCNEPMQCLCETNNGGOLCDKDLNYC 300
241 KHGSKCLPGDCRCQYQGQGLYCDKCIPIHPGCVHGINCNEPMQCLCETNNGGOLCDKDLNYC 300
301 GTHQPCPLNGGTCNTGTPDKYQCSCEGYSGPNCIEAIAHACLSDPCHNRGSCKETSLGPEC 360
301 GTHQPCPLNGGTCNTGTPDKYQCSCEGYSGPNCIEAIAHACLSDPCHNRGSCKETSLGPEC 360
361 ECSPGWTGPTCTNIDDCSPNNCSHGCTCDLVNGFKVCVPPQWTGKTQCLDANECEAKP 420
361 ECSPGWTGPTCTNIDDCSPNNCSHGCTCDLVNGFKVCVPPQWTGKTQCLDANECEAKP 420
421 CVNAKSKNLIASYCDCLPGWNGQNDININDCLGOQNDASCRDLVNGYRITCPPEYA 480
421 CVNAKSKNLIASYCDCLPGWNGQNDININDCLGOQNDASCRDLVNGYRITCPPEYA 480
481 GDHCERDIDECASNPLCNGGHCQNEINRFOLCPTGFGNLCQLDIDYCEPNPCQNGAQC 540
481 GDHCERDIDECASNPLCNGGHCQNEINRFOLCPTGFGNLCQLDIDYCEPNPCQNGAQC 540
541 YNRASDYFCKPEDYEGKNCNHLKDHCRTPPEVIDCTVAMASNDTPEGVRYISSNVCG 600
541 YNRASDYFCKPEDYEGKNCNHLKDHCRTPPEVIDCTVAMASNDTPEGVRYISSNVCG 600
601 PHGCKKSQSGKFTCDCKNGGFTGTCHENINDCESNPCRNGGTCIDGVNSYKICSDGWE 660
601 PHGCKKSQSGKFTCDCKNGGFTGTCHENINDCESNPCRNGGTCIDGVNSYKICSDGWE 660
661 GAYCETNINDCSQNPCHNGGTCRDLVNDYFCDCKNGKWKTKCHSRDSCQDEATCNGGTC 720
661 GAYCETNINDCSQNPCHNGGTCRDLVNDYFCDCKNGKWKTKCHSRDSCQDEATCNGGTC 720
721 YDEGDAFKMCPGCGWEGTTCNIARNSSCLPNPCHNGGTCVVGSGESFTVCCKEGWEGPICA 780
721 YDEGDAFKMCPGCGWEGTTCNIARNSSCLPNPCHNGGTCVVGSGESFTVCCKEGWEGPICA 780
781 QNTNDCSHPHCYNSTGTCVGDGNWYRCECAPGAGPDCRININECQSSPCAFGATCVDDEIN 840
781 QNTNDCSHPHCYNSTGTCVGDGNWYRCECAPGAGPDCRININECQSSPCAFGATCVDDEIN 840
841 GYRCVCPGSHGAKCQVSGRPCTITMGSVIPDGAKWDDDCNTCOCLNGRIACSKVWCGR 900
841 GYRCVCPGSHGAKCQVSGRPCTITMGSVIPDGAKWDDDCNTCOCLNGRIACSKVWCGR 900
901 PCLLHKHSHSCPSQSCIPILDDQCFVHPCTGVGECRSSLSLPVKTKTSDSYQDNCA 960
901 PCLLHKHSHSCPSQSCIPILDDQCFVHPCTGVGECRSSLSLPVKTKTSDSYQDNCA 960
961 ITFTFNKEMMSPGLTTEHICSELRLNMLKNVSAEYSIYIACEPSPSANNHVAISAED 1020
961 ITFTFNKEMMSPGLTTEHICSELRLNMLKNVSAEYSIYIACEPSPSANNHVAISAED 1020
1021 IRDGNPIKEITDKIIDLVSKRDNSSSLIAAFAEVRVQRPLKNRTD 1067
1021 IRDGNPIKEITDKIIDLVSKRDNSSSLIAAFAEVRVQRPLKNRTD 1067

RESULT 4

US-09-566-047-2
; Sequence 2, Application US/09566047
; Patent No. 6703198
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; Hood, Leroy
; Krantz, Ian D.
; Spinner, Nancy B.
; TITLE OF INVENTION: Methods of Diagnosing Alagille Syndrome
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego

; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/566,047
; FILING DATE: 05-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/882,046
; FILING DATE: 25-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UW 4164
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (858) 535-9001
; TELEFAX: (858) 535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1218 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-566-047-2

Query Match 100.0%; Score 6248; DB 2; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1067; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRSPTGRSGRPLSLALLCALRAKVCAGSGOFLEILSMQNVGELONGNCCGARN 60
Db 1 MRSPTGRSGRPLSLALLCALRAKVCAGSGOFLEILSMQNVGELONGNCCGARN 60

Qy 61 PGDRKTRDECDTYFKVCLKEYQSRVTAGGPCSGSGSTPVIGNTFNLKASRGNDRRI 120
Db 61 PGDRKTRDECDTYFKVCLKEYQSRVTAGGPCSGSGSTPVIGNTFNLKASRGNDRRI 120

Qy 121 VLPSPFAWPRSYTLVLEAWSSNDTVQDPSIIIEKASHGMINPSRQWTLKQNTGVAHFE 180
Db 121 VLPSPFAWPRSYTLVLEAWSSNDTVQDPSIIIEKASHGMINPSRQWTLKQNTGVAHFE 180

Qy 181 YQIRVTCDYYVYFGCNKFCRPRDDFGHYACDQNGNKTCEGWGPECNRAICRQCGSP 240
Db 181 YQIRVTCDYYVYFGCNKFCRPRDDFGHYACDQNGNKTCEGWGPECNRAICRQCGSP 240

Qy 241 KHGSKCLPGDCRCQYGMQGLYCDKCIPIHPGCVHIGICNEPWCLCETNMGWGLCDKDLNYC 300
Db 241 KHGSKCLPGDCRCQYGMQGLYCDKCIPIHPGCVHIGICNEPWCLCETNMGWGLCDKDLNYC 300

Qy 301 GTHQPLNGGTCSTNTPDKYQCSPEGYSGNCEIAEHACLSDPCHNRGSKETSLGFEC 360
Db 301 GTHQPLNGGTCSTNTPDKYQCSPEGYSGNCEIAEHACLSDPCHNRGSKETSLGFEC 360

Qy 361 ECSPGWGTGTCSTNIDDCSPNCSHGTCQDLVNGFKVCPPQWTGKTCQLDANECAKP 420
Db 361 ECSPGWGTGTCSTNIDDCSPNCSHGTCQDLVNGFKVCPPQWTGKTCQLDANECAKP 420

Qy 421 CVNAKSKNLIASYCDCLPGMWGQNCIDINDCLGQCONDASCRDLVNGYRCICPPGYA 480
Db 421 CVNAKSKNLIASYCDCLPGMWGQNCIDINDCLGQCONDASCRDLVNGYRCICPPGYA 480

Qy 481 GDHCERDIDEASNPCLNGHCONEINRFQCLCPTGFSGNCICOLDIDYCEPNPQNGAQC 540
Db 481 GDHCERDIDEASNPCLNGHCONEINRFQCLCPTGFSGNCICOLDIDYCEPNPQNGAQC 540

Qy 541 YNRASDYFCKPEDYEGKNCSHLKHCRTPCEVIDSCVTAMASNDTPEGVRYISSNVCG 600

Db 541 YNRASDYFCKPEDYEGKNCSHLKHCRTPCEVIDSCVTAMASNDTPEGVRYISSNVCG 600
Qy 601 PHGCKXSQSGGKFTCDNKGFTGTYCHENINDCESNPCRNGGTCIDGVNSYKICSDGWE 660
Db 601 PHGCKXSQSGGKFTCDNKGFTGTYCHENINDCESNPCRNGGTCIDGVNSYKICSDGWE 660
Qy 661 GAYCETNINDCSQNPCHNGGTCRDLVNDFYCDCKNGWKGTCHSRDSQCDDEATCNGGTC 720
Db 661 GAYCETNINDCSQNPCHNGGTCRDLVNDFYCDCKNGWKGTCHSRDSQCDDEATCNGGTC 720
Qy 721 YDEGDAFKMCPGGWEGTTCNIARNSSCLPNPCHNGGTCVNVGESFTVCCKEGWEGPICA 780
Db 721 YDEGDAFKMCPGGWEGTTCNIARNSSCLPNPCHNGGTCVNVGESFTVCCKEGWEGPICA 780
Qy 781 QNTNDCSPHPCYNSGTCVDGDNWYRCCECAPGAGPCDRIININECOSSPCAFGATCVDEIN 840
Db 781 QNTNDCSPHPCYNSGTCVDGDNWYRCCECAPGAGPCDRIININECOSSPCAFGATCVDEIN 840
Qy 841 GYRCVCPPGHSGAKCOEVSGRPCITWGSVIPDGAKWDDDCNTCQCLNGRIACSKVWCGPR 900
Db 841 GYRCVCPPGHSGAKCOEVSGRPCITWGSVIPDGAKWDDDCNTCQCLNGRIACSKVWCGPR 900
Qy 901 PCLLKHGSHSECPGSGQSCIPILDDQCFVHPCTGVGECRSSLSLPVTKTCTSDSYQDNCAN 960
Db 901 PCLLKHGSHSECPGSGQSCIPILDDQCFVHPCTGVGECRSSLSLPVTKTCTSDSYQDNCAN 960
Qy 961 ITFTFNKEMSPGLTTEHICSELNLIKNVAEYSIYIACBPSPANNEIHVAISAED 1020
Db 961 ITFTFNKEMSPGLTTEHICSELNLIKNVAEYSIYIACBPSPANNEIHVAISAED 1020
Qy 1021 IRDGNPKEITDKIIDLVSKRDGNSSLIAAAVAEVRVQRPLKNRTD 1067
Db 1021 IRDGNPKEITDKIIDLVSKRDGNSSLIAAAVAEVRVQRPLKNRTD 1067

RESULT 5

US-09-579-536C-1

; Sequence 1, Application US/09579536C
; Patent No. 6716974
; GENERAL INFORMATION:
; APPLICANT: MACIAG, Thomas
; APPLICANT: ZIMRIN, Ann
; APPLICANT: SMALL, Deena
; APPLICANT: PRUDOVSKY, Igor
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC METHODS AND COMPOSITIONS BASED ON JAGG
; FILE OF INVENTION: PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: 053689-5002-01
; CURRENT APPLICATION NUMBER: US/09/579,536C
; CURRENT FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/199,865
; PRIOR FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: PCT/US97/09407
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/018,841
; PRIOR FILING DATE: 1996-05-31
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 1218
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-579-536C-1

Query Match 100.0%; Score 6248; DB 2; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1067; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRSPTGRSGRPLSLALLCALRAKVCAGSGOFLEILSMQNVGELONGNCCGARN 60
Db 1 MRSPTGRSGRPLSLALLCALRAKVCAGSGOFLEILSMQNVGELONGNCCGARN 60

Qy 61 PGDRKTRDECDTYFKVCLKEYQSRVTAGGPCSGSGSTPVIGNTFNLKASRGNDRRI 120


```
Db 61 PGRKCTRDECDTYFKVCLKEYQSRVTAGGPCSFGSGSTPVIGNTFNLKASRGNDRNI 120
Qy 121 VLPFSFAPRPSYTLLEAVDSSNDTVQPSDIIIEKASHSGMINFNRQWTLKQNTGVVAHFE 180
Db 121 VLPFSFAPRPSYTLLEAVDSSNDTVQPSDIIIEKASHSGMINFNRQWTLKQNTGVVAHFE 180
Qy 181 YQIRVTCDDYYFGCNKFCRPRDDFFGHYACDQNGKTCMEGWMPGECNRAICROGCSP 240
Db 181 YQIRVTCDDYYFGCNKFCRPRDDFFGHYACDQNGKTCMEGWMPGECNRAICROGCSP 240
Qy 241 KHGSKCLPGBRCQYQWQGLYCDKCIHPGCVHGINENPQCLCETNNGQQLCDKDLNYC 300
Db 241 KHGSKCLPGBRCQYQWQGLYCDKCIHPGCVHGINENPQCLCETNNGQQLCDKDLNYC 300
Qy 301 GTHQPLNGGTCNTGPDKYQCSCEPGYSGPNCEIAEHAACLSDPCHNRGSKETSGLFEC 360
Db 301 GTHQPLNGGTCNTGPDKYQCSCEPGYSGPNCEIAEHAACLSDPCHNRGSKETSGLFEC 360
Qy 361 ECSPGWTGPTCSTNIDDCSPNNCSHGTCQDLVNGFKVCVCPQWTGKTCDLANECEAKP 420
Db 361 ECSPGWTGPTCSTNIDDCSPNNCSHGTCQDLVNGFKVCVCPQWTGKTCDLANECEAKP 420
Qy 421 CVNAKSKNLIASYYCDCLPGWQNGCDININDCLGOQNDASCRDLVNGYRCICPPGYA 480
Db 421 CVNAKSKNLIASYYCDCLPGWQNGCDININDCLGOQNDASCRDLVNGYRCICPPGYA 480
Qy 481 GDHCERDIDECASNPLCNGHCHQNEINRFQCLCPTGFSGNLCQDIDYCEPNPQNGAQC 540
Db 481 GDHCERDIDECASNPLCNGHCHQNEINRFQCLCPTGFSGNLCQDIDYCEPNPQNGAQC 540
Qy 541 YNRASDYFKCPEDYEGKNCCHLKHCRTPPCVIDSCVTAMASNDTPEGVRVYISSNVCG 600
Db 541 YNRASDYFKCPEDYEGKNCCHLKHCRTPPCVIDSCVTAMASNDTPEGVRVYISSNVCG 600
Qy 601 PHGCKKSQSGKFTCDCKNGFTGTCHENINDCESNPRNGGTCIDGVNSYKICSDGWE 660
Db 601 PHGCKKSQSGKFTCDCKNGFTGTCHENINDCESNPRNGGTCIDGVNSYKICSDGWE 660
Qy 661 GAYCETNINDCSQNPCHNGGTCDRLVNDFYCDCKNGWKGTCHSRDSQCDDEATCNGGTC 720
Db 661 GAYCETNINDCSQNPCHNGGTCDRLVNDFYCDCKNGWKGTCHSRDSQCDDEATCNGGTC 720
Qy 721 YDEGDAFKMCPGCGWGTTCNIARNSSCLPNPCHNGGTCTVNGESFTCVCKEGWGPICA 780
Db 721 YDEGDAFKMCPGCGWGTTCNIARNSSCLPNPCHNGGTCTVNGESFTCVCKEGWGPICA 780
Qy 781 QNTNDCSPHPCTYNSGTCTVDGDNMYRCECAPGAPDCRININESCQSPCAPFATCYDEIN 840
Db 781 QNTNDCSPHPCTYNSGTCTVDGDNMYRCECAPGAPDCRININESCQSPCAPFATCYDEIN 840
Qy 841 GYRCVCPGHSQAKCOEVSGRPCITWGSVTPDGAQWDDCNTCCQLNGRIACS KVMCGPR 900
Db 841 GYRCVCPGHSQAKCOEVSGRPCITWGSVTPDGAQWDDCNTCCQLNGRIACS KVMCGPR 900
Qy 901 PCLLHGHSECPGQSCIPILDDQCFVHPTGTGVECRSSSLQPVTKTSDSYQNCAN 960
Db 901 PCLLHGHSECPGQSCIPILDDQCFVHPTGTGVECRSSSLQPVTKTSDSYQNCAN 960
Qy 961 ITFTFNKEMSPGLTTEHICSELRNLNLIKNSAEYSIYIACEPSPSANNEIHVAISAED 1020
Db 961 ITFTFNKEMSPGLTTEHICSELRNLNLIKNSAEYSIYIACEPSPSANNEIHVAISAED 1020
Qy 1021 IRDDGNPIKEITDKIIDLVSKRDNSSLIJAAVAEVRVQRPLKNRTD 1067
Db 1021 IRDDGNPIKEITDKIIDLVSKRDNSSLIJAAVAEVRVQRPLKNRTD 1067
```

RESULT 6

US-09-949-016-5902
; Sequence 5902, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5902
; LENGTH: 1218
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-5902

Query Match 100.0%; Score 6248; DB 2; Length 1218;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1067; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 MRSPTRGSRGRLSLLALLCALRAKVCASGQFELEILSMQNVGELQNGCCGARN 60
Db 1 MRSPTRGSRGRLSLLALLCALRAKVCASGQFELEILSMQNVGELQNGCCGARN 60
Qy 61 PGRKCTRDECDTYFKVCLKEYQSRVTAGGPCSFGSGSTPVIGNTFNLKASRGNDRNI 120
Db 61 PGRKCTRDECDTYFKVCLKEYQSRVTAGGPCSFGSGSTPVIGNTFNLKASRGNDRNI 120
Qy 121 VLPFSFAPRPSYTLLEAVDSSNDTVQPSDIIIEKASHSGMINFNRQWTLKQNTGVVAHFE 180
Db 121 VLPFSFAPRPSYTLLEAVDSSNDTVQPSDIIIEKASHSGMINFNRQWTLKQNTGVVAHFE 180
Qy 181 YQIRVTCDDYYFGCNKFCRPRDDFFGHYACDQNGKTCMEGWMPGECNRAICROGCSP 240
Db 181 YQIRVTCDDYYFGCNKFCRPRDDFFGHYACDQNGKTCMEGWMPGECNRAICROGCSP 240
Qy 241 KHGSKCLPGBRCQYQWQGLYCDKCIHPGCVHGINENPQCLCETNNGQQLCDKDLNYC 300
Db 241 KHGSKCLPGBRCQYQWQGLYCDKCIHPGCVHGINENPQCLCETNNGQQLCDKDLNYC 300
Qy 301 GTHQPLNGGTCNTGPDKYQCSCEPGYSGPNCEIAEHAACLSDPCHNRGSKETSGLFEC 360
Db 301 GTHQPLNGGTCNTGPDKYQCSCEPGYSGPNCEIAEHAACLSDPCHNRGSKETSGLFEC 360
Qy 361 ECSPGWTGPTCSTNIDDCSPNNCSHGTCQDLVNGFKVCVCPQWTGKTCDLANECEAKP 420
Db 361 ECSPGWTGPTCSTNIDDCSPNNCSHGTCQDLVNGFKVCVCPQWTGKTCDLANECEAKP 420
Qy 421 CVNAKSKNLIASYYCDCLPGWQNGCDININDCLGOQNDASCRDLVNGYRCICPPGYA 480
Db 421 CVNAKSKNLIASYYCDCLPGWQNGCDININDCLGOQNDASCRDLVNGYRCICPPGYA 480
Qy 481 GDHCERDIDECASNPLCNGHCHQNEINRFQCLCPTGFSGNLCQDIDYCEPNPQNGAQC 540
Db 481 GDHCERDIDECASNPLCNGHCHQNEINRFQCLCPTGFSGNLCQDIDYCEPNPQNGAQC 540
Qy 541 YNRASDYFKCPEDYEGKNCCHLKHCRTPPCVIDSCVTAMASNDTPEGVRVYISSNVCG 600
Db 541 YNRASDYFKCPEDYEGKNCCHLKHCRTPPCVIDSCVTAMASNDTPEGVRVYISSNVCG 600
Qy 601 PHGCKKSQSGKFTCDCKNGFTGTCHENINDCESNPRNGGTCIDGVNSYKICSDGWE 660
Db 601 PHGCKKSQSGKFTCDCKNGFTGTCHENINDCESNPRNGGTCIDGVNSYKICSDGWE 660
Qy 661 GAYCETNINDCSQNPCHNGGTCDRLVNDFYCDCKNGWKGTCHSRDSQCDDEATCNGGTC 720
Db 661 GAYCETNINDCSQNPCHNGGTCDRLVNDFYCDCKNGWKGTCHSRDSQCDDEATCNGGTC 720
Qy 721 YDEGDAFKMCPGCGWGTTCNIARNSSCLPNPCHNGGTCTVNGESFTCVCKEGWGPICA 780
Db 721 YDEGDAFKMCPGCGWGTTCNIARNSSCLPNPCHNGGTCTVNGESFTCVCKEGWGPICA 780
```

```
Db 721 YDEGDAFKCMPCGGWEGTTCTNIARNSSCLPNPCHNGGTCVVGNGESFTVCCKEGWGPICA 780
Qy 781 QNTNDCSPHPCYNSGTCVGDGNNWYRCECAPGFAGPDCRININBCQSSPCAFGATCVDEIN 840
Db 781 QNTNDCSPHPCYNSGTCVGDGNNWYRCECAPGFAGPDCRININBCQSSPCAFGATCVDEIN 840
Qy 841 GYRCVCPGGHSGAKQCVSRGSPCITWGSVIPDGAkWDDCNTCQCLNGRIACSKVWCGPR 900
Db 841 GYRCVCPGGHSGAKQCVSRGSPCITWGSVIPDGAkWDDCNTCQCLNGRIACSKVWCGPR 900
Qy 901 PCLLHKHSGPCSPGQSCIPILDDQCFVHPCTGVGECRSSSLQPVKTKTSDSYQDNCA 960
Db 901 PCLLHKHSGPCSPGQSCIPILDDQCFVHPCTGVGECRSSSLQPVKTKTSDSYQDNCA 960
Qy 961 ITFTFNKEMMSPLTTEHICSELRLNLIKVNLSAEYSIYIACPPSPANNEIHVAISAED 1020
Db 961 ITFTFNKEMMSPLTTEHICSELRLNLIKVNLSAEYSIYIACPPSPANNEIHVAISAED 1020
Qy 1021 IRDDGNPIKEITDKIIDLVSKRDNSSSLIAAFAEVRVQRRLKNRTD 1067
Db 1021 IRDDGNPIKEITDKIIDLVSKRDNSSSLIAAFAEVRVQRRLKNRTD 1067

RESULT 7
US-09-949-016-10297
; Sequence 10297, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10297
; LENGTH: 1254
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10297

Query Match 100.0%; Score 6248; DB 2; Length 1254;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1067; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRSPTRGSRPLSLIALCALRAKVCAGSGQFEILSMQNVNGLQNGNCCGGARN 60
Db 37 MRSPTRGSRPLSLIALCALRAKVCAGSGQFEILSMQNVNGLQNGNCCGGARN 96
Qy 61 PGDRKTRDECDTFFKVCLEKYQSRVTAGGCPGSGSTPVIGNTFNLKASRGNDNRNI 120
Db 97 PGDRKTRDECDTFFKVCLEKYQSRVTAGGCPGSGSTPVIGNTFNLKASRGNDNRNI 156
Qy 121 VLPSFAWPRSVTLLEAWDSSNDTVQDPSIIIEKASHGSMINPSRQWTLKQNTGVAFHE 180
Db 157 VLPSFAWPRSVTLLEAWDSSNDTVQDPSIIIEKASHGSMINPSRQWTLKQNTGVAFHE 216
Qy 181 YQIRVTCDDYYFGCNKFCRPRDDFFGHYACDQNGNKTCHMEGWPCECNRAICRQGSP 240
Db 217 YQIRVTCDDYYFGCNKFCRPRDDFFGHYACDQNGNKTCHMEGWPCECNRAICRQGSP 276
Qy 241 KHGSKLPGDCRCQYWGQGLYCDKICIPHPGCVHGI CNBPWQCLCETNMGWGLCDKDLNYC 300
Db 277 KHGSKLPGDCRCQYWGQGLYCDKICIPHPGCVHGI CNBPWQCLCETNMGWGLCDKDLNYC 336
Qy 301 GTHQPCNLGGTCSNTGPDKYQCSCEPGYSGPNCETAEHACLSDPCPNRSGKETSLSGFEC 360
```

```
Db 337 GTHQPCNLGGTCSNTGPDKYQCSCEPGYSGPNCETAEHACLSDPCPNRSGKETSLSGFEC 396
Qy 361 ECSPGWGTGTCSTNIDDCSPNNCSHGCTCODLVNGFKVCYCPQWTKTCTOLDANECEAKP 420
Db 397 ECSPGWGTGTCSTNIDDCSPNNCSHGCTCODLVNGFKVCYCPQWTKTCTOLDANECEAKP 456
Qy 421 CYNKASCKNLILASYCDCLPGMMQNCNININDCLGOQNDASCRDLVNGYRNICPPGYA 480
Db 457 CYNKASCKNLILASYCDCLPGMMQNCNININDCLGOQNDASCRDLVNGYRNICPPGYA 516
Qy 481 GDCERDIDECASNPCNLGGHCHQNEINRFQCLCPTGFSNLCOLDIDYCEPNPQNGAQC 540
Db 517 GDCERDIDECASNPCNLGGHCHQNEINRFQCLCPTGFSNLCOLDIDYCEPNPQNGAQC 576
Qy 541 YNRASDYFKCPBPDYEGKNCCHLKHCRTPPCBVIDSCIVAMASNDTPBGRVYISSNVC 600
Db 577 YNRASDYFKCPBPDYEGKNCCHLKHCRTPPCBVIDSCIVAMASNDTPBGRVYISSNVC 636
Qy 601 PHGCKKSQSGGKFTCDCKNGFTCTYCHENINDCESNPRNGGTCIDGVNSYKICISDGWE 660
Db 637 PHGCKKSQSGGKFTCDCKNGFTCTYCHENINDCESNPRNGGTCIDGVNSYKICISDGWE 696
Qy 661 GAYCETNINDCSQNPCHNGGTCRDVLVNDFYCDCKNGWKGTCHSRDSQCDDEATCNNGGTC 720
Db 697 GAYCETNINDCSQNPCHNGGTCRDVLVNDFYCDCKNGWKGTCHSRDSQCDDEATCNNGGTC 756
Qy 721 YDEGDAFKCMPCGGWEGTTCTNIARNSSCLPNPCHNGGTCVVGNGESFTVCCKEGWGPICA 780
Db 757 YDEGDAFKCMPCGGWEGTTCTNIARNSSCLPNPCHNGGTCVVGNGESFTVCCKEGWGPICA 816
Qy 781 QNTNDCSPHPCYNSGTCVGDGNNWYRCECAPGFAGPDCRININBCQSSPCAFGATCVDEIN 840
Db 817 QNTNDCSPHPCYNSGTCVGDGNNWYRCECAPGFAGPDCRININBCQSSPCAFGATCVDEIN 876
Qy 841 GYRCVCPGGHSGAKQCVSRGSPCITWGSVIPDGAkWDDCNTCQCLNGRIACSKVWCGPR 900
Db 877 GYRCVCPGGHSGAKQCVSRGSPCITWGSVIPDGAkWDDCNTCQCLNGRIACSKVWCGPR 936
Qy 901 PCLLHKHSGPCSPGQSCIPILDDQCFVHPCTGVGECRSSSLQPVKTKTSDSYQDNCA 960
Db 937 PCLLHKHSGPCSPGQSCIPILDDQCFVHPCTGVGECRSSSLQPVKTKTSDSYQDNCA 996
Qy 961 ITFTFNKEMMSPLTTEHICSELRLNLIKVNLSAEYSIYIACPPSPANNEIHVAISAED 1020
Db 997 ITFTFNKEMMSPLTTEHICSELRLNLIKVNLSAEYSIYIACPPSPANNEIHVAISAED 1056
Qy 1021 IRDDGNPIKEITDKIIDLVSKRDNSSSLIAAFAEVRVQRRLKNRTD 1067
Db 1057 IRDDGNPIKEITDKIIDLVSKRDNSSSLIAAFAEVRVQRRLKNRTD 1103
```

```
RESULT 8
US-09-917-254-85
; Sequence 85, Application US/09917254
; Patent No. 6703204
; GENERAL INFORMATION:
; APPLICANT: Mutter, George
; APPLICANT: Baak, Jan
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer
; FILE REFERENCE: B0801/7224(JRV)
; CURRENT APPLICATION NUMBER: US/09/917,254
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/222,093
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 85
; LENGTH: 1218
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-917-254-85
```



```
Db 241 KHGSKLPGDCRCQYWGQGLYCDKCIHPGCVHGCNPNWQCLCEITWGGQLCDKOLNYC 300
Qy 301 GTHQPLNGGTCSTNTGDPDKYQCSCEGYSNPNCIEAHACLSDPCHNRGSKETSIGFEC 360
Db 301 GTHQPLNGGTCSTNTGDPDKYQCSCEGYSNPNCIEAHACLSDPCHNRGSKETSIGFEC 360
Qy 361 ECSPGWTGPTCSTNIDDCSPNNCSHGCTCDLVNGFKVCVPPQWTGKTCDLDANECEAKP 420
Db 361 ECSPGWTGPTCSTNIDDCSPNNCSHGCTCDLVNGFKVCVPPQWTGKTCDLDANECEAKP 420
Qy 421 CVNAKSKNLIASYCDCLPGWQGNCDININDCLGQCNDAASCRDLVNGYRCICPPGYA 480
Db 421 CVNAKSKNLIASYCDCLPGWQGNCDININDCLGQCNDAASCRDLVNGYRCICPPGYA 480
Qy 481 GDHCERDIDECASNPLNGHGCQNEINRFQCLCPTGFGSGLCQDLDIDYCEPNPCQNGAQC 540
Db 481 GDHCERDIDECASNPLNGHGCQNEINRFQCLCPTGFGSGLCQDLDIDYCEPNPCQNGAQC 540
Qy 541 YNRASDYFCCKPEDYEGKNCSHLKHCHRTTTPCEVIDSCTVMAASNDTPEGVRYISSNVCG 600
Db 541 YNRASDYFCCKPEDYEGKNCSHLKHCHRTTTPCEVIDSCTVMAASNDTPEGVRYISSNVCG 600
Qy 601 PHGCKKSQSGKFTCDCKNGFTGTGTYCHENINDCESNFCRNGGTCIDGVNSYKICSDGWE 660
Db 601 PHGCKKSQSGKFTCDCKNGFTGTGTYCHENINDCESNFCRNGGTCIDGVNSYKICSDGWE 660
Qy 661 GAYCETNINDCSQNPCHNGGTCRLVNDFYCDCKNGWKGTCHSRDSQCDDEATCNGGTC 720
Db 661 GAYCETNINDCSQNPCHNGGTCRLVNDFYCDCKNGWKGTCHSRDSQCDDEATCNGGTC 720
Qy 721 YDEGDAFKMCPGWEGTTCNIARNSCLNPNCHNGGTCVVGNSFTVCCKEGWGPICA 780
Db 721 YDEGDAFKMCPGWEGTTCNIARNSCLNPNCHNGGTCVVGNSFTVCCKEGWGPICA 780
Qy 781 QNTNDCSPHPCYNSGTCVGDGNWYRCBAPGAFGDCRININEQSSPCAFAGTCVDEIN 840
Db 781 QNTNDCSPHPCYNSGTCVGDGNWYRCBAPGAFGDCRININEQSSPCAFAGTCVDEIN 840
Qy 841 GYRCVCPGHSKAKQEVSGRPCITWGSVIPDGAKWDDDCNTCQCLNGRIACSKVWCGR 900
Db 841 GYRCVCPGHSKAKQEVSGRPCITWGSVIPDGAKWDDDCNTCQCLNGRIACSKVWCGR 900
Qy 901 PCLLKHGSECPQSQCIPILDDQCFVHPCTGVGECRSSLSLPVKTKTSDSYQDNCAN 960
Db 901 PCLLKHGSECPQSQCIPILDDQCFVHPCTGVGECRSSLSLPVKTKTSDSYQDNCAN 960
Qy 961 ITFTFNKEMMSPGLTTEHICSELRLNLIKNSVAEYSIYIACEPSPSANNIEHVAISAED 1020
Db 961 ITFTFNKEMMSPGLTTEHICSELRLNLIKNSVAEYSIYIACEPSPSANNIEHVAISAED 1020
Qy 1021 IRDGNPIKETDKIIDLVRKDGNSLLIAAFAEVRVQRRLKQRTD 1067
Db 1021 IRDGNPIKETDKIIDLVRKDGNSLLIAAFAEVRVQRRLKQRTD 1067
```

RESULT 10

```
US-08-611-729A-6
; Sequence 6, Application US/08611729A
; Patent No. 6004924
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowitz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Teakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; SERRATE GENE AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,729A
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1218 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-611-729A-6
```

```
Query Match 99.8%; Score 6233; DB 2; Length 1218;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1064; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MRSPTRGSRGRPLSLALLCALRAKVCAGSGQFEILEILSMQNVNGLQNGNCCGARN 60
Db 1 MRSPTRGSRGRPLSLALLCALRAKVCAGSGQFEILEILSMQNVNGLQNGNCCGARN 60
Qy 61 PGRKCTRDCEDTYFKVCLKEYOSRVTAGGPGSGSTPVIIGNTFNKASRNDNRNI 120
Db 61 PGRKCTRDCEDTYFKVCLKEYOSRVTAGGPGSGSTPVIIGNTFNKASRNDNRNI 120
Qy 121 VLPFSFAPRPSYTLLEAVDSSNDTVQDSIIIEKASHSGMINFSRQWTLKQNTGVAHFE 180
Db 121 VLPFSFAPRPSYTLLEAVDSSNDTVQDSIIIEKASHSGMINFSRQWTLKQNTGVAHFE 180
Qy 181 YQIRVTCDDYYYGFCNKFCRPRDDPFPHVACDQNGNKTQMEGWMGPECNRAICROGCSP 240
Db 181 YQIRVTCDDYYYGFCNKFCRPRDDPFPHVACDQNGNKTQMEGWMGPECNRAICROGCSP 240
Qy 241 KHGSKLPGDCRCQYWGQGLYCDKCIHPGCVHGCNPNWQCLCEITWGGQLCDKOLNYC 300
Db 241 KHGSKLPGDCRCQYWGQGLYCDKCIHPGCVHGCNPNWQCLCEITWGGQLCDKOLNYC 300
Qy 301 GTHQPLNGGTCSTNTGDPDKYQCSCEGYSNPNCIEAHACLSDPCHNRGSKETSIGFEC 360
Db 301 GTHQPLNGGTCSTNTGDPDKYQCSCEGYSNPNCIEAHACLSDPCHNRGSKETSIGFEC 360
Qy 361 ECSPGWTGPTCSTNIDDCSPNNCSHGCTCDLVNGFKVCVPPQWTGKTCDLDANECEAKP 420
Db 361 ECSPGWTGPTCSTNIDDCSPNNCSHGCTCDLVNGFKVCVPPQWTGKTCDLDANECEAKP 420
Qy 421 CVNAKSKNLIASYCDCLPGWQGNCDININDCLGQCNDAASCRDLVNGYRCICPPGYA 480
Db 421 CVNAKSKNLIASYCDCLPGWQGNCDININDCLGQCNDAASCRDLVNGYRCICPPGYA 480
Qy 481 GDHCERDIDECASNPLNGHGCQNEINRFQCLCPTGFGSGLCQDLDIDYCEPNPCQNGAQC 540
Db 481 GDHCERDIDECASNPLNGHGCQNEINRFQCLCPTGFGSGLCQDLDIDYCEPNPCQNGAQC 540
Qy 541 YNRASDYFCCKPEDYEGKNCSHLKHCHRTTTPCEVIDSCTVMAASNDTPEGVRYISSNVCG 600
Db 541 YNRASDYFCCKPEDYEGKNCSHLKHCHRTTTPCEVIDSCTVMAASNDTPEGVRYISSNVCG 600
```

Db 541 YNRADYFCKPEDYEGKNCNHLKHCHRTTCEVIDSCTVAMASNDTPGVRYISSLNVC 600
Qy 601 PHGCKSQSGKFTCDNKGFTGTGTYCHENINDCESNPCRNGGTCTIDGVNSYKICISDGWE 660
Db 601 PHGCKSQSGKFTCDNKGFTGTGTYCHENINDCESNPCRNGGTCTIDGVNSYKICISDGWE 660
Qy 661 GAYCETINDCSONPCHNGGTCRDVLVNDYFCDCKNGWKGTCHSRDSCQDEATCNGGTC 720
Db 661 GAYCETINDCSONPCHNGGTCRDVLVNDYFCDCKNGWKGTCHSRDSCQDEATCNGGTC 720
Qy 721 YDEGDAFKMCPGWEFTTCNIARNSSCLPNPCHNGGTCVNVGSEFTCVCKEGWGPICA 780
Db 721 YDEGDAFKMCPGWEFTTCNIARNSSCLPNPCHNGGTCVNVGSEFTCVCKEGWGPICA 780
Qy 781 QNTNDCSPHPCYNSGTCTVDGNWYRCBAPGAGPDCRININEQSSPCAFGATCVDEIN 840
Db 781 QNTNDCSPHPCYNSGTCTVDGNWYRCBAPGAGPDCRININEQSSPCAFGATCVDEIN 840
Qy 841 GYRCVCPGHSKACQEVSGRPCTITMGSVIPDGAKWDDDCNTCQCLNGRIACSKVWCGPR 900
Db 841 GYRCVCPGHSKACQEVSGRPCTITMGSVIPDGAKWDDDCNTCQCLNGRIACSKVWCGPR 900
Qy 901 PCLLKHGSECPGSCIPILDDQCFVHPCTGVGECRSSILOPVKTKTSDSYIYQDNAN 960
Db 901 PCLLKHGSECPGSCIPILDDQCFVHPCTGVGECRSSILOPVKTKTSDSYIYQDNAN 960
Qy 961 ITFTFNKEMSPGLTTEHICSELANLILKNVSAEYIYIACBSPSANNIEHVAISAED 1020
Db 961 ITFTFNKEMSPGLTTEHICSELANLILKNVSAEYIYIACBSPSANNIEHVAISAED 1020
Qy 1021 IRDGNPIKEITDKIIDLVSXRDGNSLLIAAEVVRQRPRLKURTD 1067
Db 1021 IRDGNPIKEITDKIIDLVSXRDGNSLLIAAEVVRQRPRLKURTD 1067

RESULT 11

US-09-195-524-6
; Sequence 6, Application US/09195524
; Patent No. 6703489
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowitz, David
; APPLICANT: Henrique, Domingos M. P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; TITLE OF INVENTION: SRRATE GENE AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/195,524
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US-08/611,729
; FILING DATE: 06-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872

; REFERENCE/DOCKET NUMBER: 7326-037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1218 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-09-195-524-6

Query Match 99.8%; Score 6233; DB 2; Length 1218;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1064; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MRSPTRGSRPLSLALLCALRAKVCASGOFELEILSMQNVNGLONGNCCGARN 60
Db 1 MRSPTRGSRPLSLALLCALRAKVCASGOFELEILSMQNVNGLONGNCCGARN 60
Qy 61 PGDRKCTRDCEDTYFKVCLKEYQSRVTAGGPCSGSGSTPVIIGNTFNLKASRNDNR 120
Db 61 PGDRKCTRDCEDTYFKVCLKEYQSRVTAGGPCSGSGSTPVIIGNTFNLKASRNDNR 120
Qy 121 VLPFSPAWPRSYTLLEAVDSSNDTVQPSDSIIIEKASHSGMINPSRQWOTLKQNTGVAHFE 180
Db 121 VLPFSPAWPRSYTLLEAVDSSNDTVQPSDSIIIEKASHSGMINPSRQWOTLKQNTGVAHFE 180
Qy 181 YQIRVTCDDYVYFGCNKFCRPRDDPFHYACDQNGKTCMEGWMPGECNRAICROGSP 240
Db 181 YQIRVTCDDYVYFGCNKFCRPRDDPFHYACDQNGKTCMEGWMPGECNRAICROGSP 240
Qy 241 KHGSKLPDRCQYQWQGLYCDKCPHPCGVHGCINPEWQCLCETNMGQOLCDKDLNYC 300
Db 241 KHGSKLPDRCQYQWQGLYCDKCPHPCGVHGCINPEWQCLCETNMGQOLCDKDLNYC 300
Qy 301 GTHQPCLINGGTCSTNTGPDKYQCSCEPGYSGPNCEIAEHAACLSDPCHNRSGCKETS LGFEC 360
Db 301 GTHQPCLINGGTCSTNTGPDKYQCSCEPGYSGPNCEIAEHAACLSDPCHNRSGCKETS LGFEC 360
Qy 361 ECSFGWTGPTCSNTNIDDCSPNCSHGTCQDLVNGFKVCVPPQWTKTCDLDANECAKP 420
Db 361 ECSFGWTGPTCSNTNIDDCSPNCSHGTCQDLVNGFKVCVPPQWTKTCDLDANECAKP 420
Qy 421 CVNAKSKNLIASVYCDCLPGWNGQCDININCLGQCDNDASCRDLVNGYRCICPFGYA 480
Db 421 CVNAKSKNLIASVYCDCLPGWNGQCDININCLGQCDNDASCRDLVNGYRCICPFGYA 480
Qy 481 GDHCERDIDECASNPLCNGGHCQNEINRFQCLCPTGFSNLCQLDIDYCEPNPCQNGAQC 540
Db 481 GDHCERDIDECASNPLCNGGHCQNEINRFQCLCPTGFSNLCQLDIDYCEPNPCQNGAQC 540
Qy 541 YNRADYFCKPEDYEGKNCNHLKHCHRTTCEVIDSCTVAMASNDTPGVRYISSLNVC 600
Db 541 YNRADYFCKPEDYEGKNCNHLKHCHRTTCEVIDSCTVAMASNDTPGVRYISSLNVC 600
Qy 601 PHGCKSQSGKFTCDNKGFTGTGTYCHENINDCESNPCRNGGTCTIDGVNSYKICISDGWE 660
Db 601 PHGCKSQSGKFTCDNKGFTGTGTYCHENINDCESNPCRNGGTCTIDGVNSYKICISDGWE 660
Qy 661 GAYCETINDCSONPCHNGGTCRDVLVNDYFCDCKNGWKGTCHSRDSCQDEATCNGGTC 720
Db 661 GAYCETINDCSONPCHNGGTCRDVLVNDYFCDCKNGWKGTCHSRDSCQDEATCNGGTC 720
Qy 721 YDEGDAFKMCPGWEFTTCNIARNSSCLPNPCHNGGTCVNVGSEFTCVCKEGWGPICA 780
Db 721 YDEGDAFKMCPGWEFTTCNIARNSSCLPNPCHNGGTCVNVGSEFTCVCKEGWGPICA 780
Qy 781 QNTNDCSPHPCYNSGTCTVDGNWYRCBAPGAGPDCRININEQSSPCAFGATCVDEIN 840
Db 781 QNTNDCSPHPCYNSGTCTVDGNWYRCBAPGAGPDCRININEQSSPCAFGATCVDEIN 840


```
Qy 841 GYRCVCPGSHGAKQEVSRPCITMGSVIPDGAkWDDDCNTCQCLNGRIACSKVWGCP 900
Db 841 GYRCVCPGSHGAKQEVSRPCITMGSVIPDGAkWDDDCNTCQCLNGRIACSKVWGCP 900
Qy 901 PCLLKHGSHSCPSQSCIPILDDQCFVHPCTGVGECRSSLQPVKTKTSDSYQDNCA 960
Db 901 PCLLKHGSHSCPSQSCIPILDDQCFVHPCTGVGECRSSLQPVKTKTSDSYQDNCA 960
Qy 961 ITFTFNKEMSPGLTTEHICSELNRLNLIKNSAEYSIYIACBSPSANNEIHVAISA 1020
Db 961 ITFTFNKEMSPGLTTEHICSELNRLNLIKNSAEYSIYIACBSPSANNEIHVAISA 1020
Qy 1021 IRDDGNPIKEITDKIIDLVRKDNSSLLIAAVERVQRRPLKRTD 1067
Db 1021 IRDDGNPIKEITDKIIDLVRKDNSSLLIAAVERVQRRPLKRTD 1067

RESULT 12
US-09-310-685-4
; Sequence 4, Application US/09310685
; Patent No. 6887475
; GENERAL INFORMATION:
; APPLICANT: Lamb, Jonathan R
; APPLICANT: Dallman, Margaret J
; APPLICANT: Hoyne, Gerard F
; TITLE OF INVENTION: No. 6887475ch
; FILE REFERENCE: 674525-2001
; CURRENT APPLICATION NUMBER: US/09/310,685
; CURRENT FILING DATE: 1999-05-04
; PRIOR APPLICATION NUMBER: GB 9623236.8
; PRIOR FILING DATE: 1996-11-07
; PRIOR APPLICATION NUMBER: GB 9715674.9
; PRIOR FILING DATE: 1997-07-24
; PRIOR APPLICATION NUMBER: GB 9719350.2
; PRIOR FILING DATE: 1997-09-11
; PRIOR APPLICATION NUMBER: PCT/GB97/03058
; PRIOR FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 1218
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-310-685-4

Query Match 99.8%; Score 6233; DB 2; Length 1218;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1064; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 MRSPTRGSRPLSLLLALLCALRAKVCAGSGQFELEILSMQNVNGBELQNGCCGGARN 60
Db 1 MRSPTRGSRPLSLLLALLCALRAKVCAGSGQFELEILSMQNVNGBELQNGCCGGARN 60
Qy 61 PGDRKCTREDCDTYFKVCLKEYQSRVTAGGCPSCFGSGSTPVIGGNTFNKASRGNDNR 120
Db 61 PGDRKCTREDCDTYFKVCLKEYQSRVTAGGCPSCFGSGSTPVIGGNTFNKASRGNDNR 120
Qy 121 VLPSFAWPRTYLLVEAWDSNDTVQDPSIIEKASHGSMINPSRQWTLKONTGVAHFE 180
Db 121 VLPSFAWPRTYLLVEAWDSNDTVQDPSIIEKASHGSMINPSRQWTLKONTGVAHFE 180
Qy 181 YOIRVTCDDYYGFGCNKFCPRDDFGHYACDQNGKNTQMEGWMGPCECRAICRQCGSP 240
Db 181 YOIRVTCDDYYGFGCNKFCPRDDFGHYACDQNGKNTQMEGWMGPCECRAICRQCGSP 240
Qy 241 KHGSKLPGDCRCQYGMWGLYCDKICIPHPGCVHGI CNEPWQCLCETNWWGGLCDKLDNYC 300
Db 241 KHGSKLPGDCRCQYGMWGLYCDKICIPHPGCVHGI CNEPWQCLCETNWWGGLCDKLDNYC 300
Qy 301 GTHQPLNGGTCSNTGPPDKYQCSPEBYSGNCEIAEHACLSDPCHNRGSKCTSLGPEC 360
Db 301 GTHQPLNGGTCSNTGPPDKYQCSPEBYSGNCEIAEHACLSDPCHNRGSKCTSLGPEC 360
```

```
Qy 361 ECSFGWTGPTCSTNIDDCSPNNCSHGCTCQDLVNGFKVCVPPQWTGKTCOLDANECEAKP 420
Db 361 ECSFGWTGPTCSTNIDDCSPNNCSHGCTCQDLVNGFKVCVPPQWTGKTCOLDANECEAKP 420
Qy 421 CVNAKCKNLIAIYYCDCLPGWNGQNCNINDCLGQCCQNDASCRDLVNGRYCICPPGYA 480
Db 421 CVNAKCKNLIAIYYCDCLPGWNGQNCNINDCLGQCCQNDASCRDLVNGRYCICPPGYA 480
Qy 481 GDHCERDIDECASNPCPLNGHCHCONEINRFQCLCPTGFSNLCOLDIDYCEPNPCQNGAQC 540
Db 481 GDHCERDIDECASNPCPLNGHCHCONEINRFQCLCPTGFSNLCOLDIDYCEPNPCQNGAQC 540
Qy 541 YNRASDYFKCPCPDYEGKNCNHLKDHCRTPPCVIDSCTVMAASNDTPEGVRYISSNVCG 600
Db 541 YNRASDYFKCPCPDYEGKNCNHLKDHCRTPPCVIDSCTVMAASNDTPEGVRYISSNVCG 600
Qy 601 PHGCKSQSGGKFTCDCKNGFTGTTCYCHENINDCESNPCRNGGTCIDGVNSYKICSDGWE 660
Db 601 PHGCKSQSGGKFTCDCKNGFTGTTCYCHENINDCESNPCRNGGTCIDGVNSYKICSDGWE 660
Qy 661 GAYCETNINDCSQNPCHNGGTCRDLVNDFYCDCKNGWKGTCHSRDSQCDDEATCNGGTC 720
Db 661 GAYCETNINDCSQNPCHNGGTCRDLVNDFYCDCKNGWKGTCHSRDSQCDDEATCNGGTC 720
Qy 721 YDEGDAFKCMCPGSGWETTCNIARNSSCLPNPCHNGGTCVVGESFTVCCKEGWGPICA 780
Db 721 YDEGDAFKCMCPGSGWETTCNIARNSSCLPNPCHNGGTCVVGESFTVCCKEGWGPICA 780
Qy 781 QNTNDCSPPHPCVNSGTCVGDNDWYRECAPGAPGDCRININECQSSPCAFGATCVDEN 840
Db 781 QNTNDCSPPHPCVNSGTCVGDNDWYRECAPGAPGDCRININECQSSPCAFGATCVDEN 840
Qy 841 GYRCVCPGSHGAKQEVSRPCITMGSVIPDGAkWDDDCNTCQCLNGRIACSKVWGCP 900
Db 841 GYRCVCPGSHGAKQEVSRPCITMGSVIPDGAkWDDDCNTCQCLNGRIACSKVWGCP 900
Qy 901 PCLLKHGSHSCPSQSCIPILDDQCFVHPCTGVGECRSSLQPVKTKTSDSYQDNCA 960
Db 901 PCLLKHGSHSCPSQSCIPILDDQCFVHPCTGVGECRSSLQPVKTKTSDSYQDNCA 960
Qy 961 ITFTFNKEMSPGLTTEHICSELNRLNLIKNSAEYSIYIACBSPSANNEIHVAISA 1020
Db 961 ITFTFNKEMSPGLTTEHICSELNRLNLIKNSAEYSIYIACBSPSANNEIHVAISA 1020
Qy 1021 IRDDGNPIKEITDKIIDLVRKDNSSLLIAAVERVQRRPLKRTD 1067
Db 1021 IRDDGNPIKEITDKIIDLVRKDNSSLLIAAVERVQRRPLKRTD 1067
```

```
RESULT 13
US-09-214-278-7
; Sequence 7, Application US/09214278
; Patent No. 6291210
; GENERAL INFORMATION:
; APPLICANT: Sakano, Seiji
; APPLICANT: Itoh, Akira
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8576
; CURRENT APPLICATION NUMBER: US/09/214,278
; CURRENT FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 1218
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-214-278-7
```

```
Query Match 99.7%; Score 6228; DB 2; Length 1218;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1065; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MRSPTRGSRPLSLLLALLCALRAKVCAGSGQFELEILSMQNVNGBELQNGCCGGARN 60
```

```
Db 1 MRSPTTRGRSRPLSLALLCALRAKVCASGQFELILSMQNVGELQNGCCGARN 60
Qy 61 PGDRKCTRDECDTYFKVCLKEYQSRVTAGGPCSGSGSTPVIGNTFNKASRGNDRRI 120
Db 61 PGDRKCTRDESDTYFKVCLKEYQSRVTAGGPCSGSGSTPVIGNTFNKASRGNDRRI 120
Qy 121 VLPSPFAPRSYTLLEAVDSSNDTVQPDSEIEKASHGMINPSPQWTLKQNTGVVAHPE 180
Db 121 VLPSPFAPRSYTLLEAVDSSNDTVQPDSEIEKASHGMINPSPQWTLKQNTGVVAHPE 180
Qy 181 YQIRVTCDDYYFGCNKFCRPRDDFFGHYACDQNGKTCMEGWMGPECNRAICRQCSP 240
Db 181 YQIRVTCDDYYFGCNKFCRPRDDFFGHYACDQNGKTCMEGWMGPECNRAICRQCSP 240
Qy 241 KHGSKCLPGDCRCQYQWGLYCDKCI PHPGCVHGINPQCLCETNMGWGLCDKDLNYC 300
Db 241 KHGSKCLPGDCRCQYQWGLYCDKCI PHPGCVHGINPQCLCETNMGWGLCDKDLNYC 300
Qy 301 GTHQPCINGTCSNTGPDKYQCSCEGYSGPNCIEAHACLSDPCHNRGSKETSLSGFEC 360
Db 301 GTHQPCINGTCSNTGPDKYQCSCEGYSGPNCIEAHACLSDPCHNRGSKETSLSGFEC 360
Qy 361 ECSPGWTGPTCSNTIDDCSPNNSHGTCQDLVNGFKVCVCPPOWTGKTCOLDANECEAKP 420
Db 361 ECSPGWTGPTCSNTIDDCSPNNSHGTCQDLVNGFKVCVCPPOWTGKTCOLDANECEAKP 420
Qy 421 CVNAKSKNLIASYCDCLPGWQONCDININDCLGQONCDASCRDLVNGYRCICPPGYA 480
Db 421 CVNAKSKNLIASYCDCLPGWQONCDININDCLGQONCDASCRDLVNGYRCICPPGYA 480
Qy 481 GDHCERDIDECASNPCLNGHCHQNEINRFQCLCPTGFSGNLCOLDIDYCEPNPCQNGAQC 540
Db 481 GDHCERDIDECASNPCLNGHCHQNEINRFQCLCPTGFSGNLCOLDIDYCEPNPCQNGAQC 540
Qy 541 YNRASDYFCCKPEDYEGKNCNHLKHCRTPPCEVIDSCVTAMASNDTPEGVYISSNVCG 600
Db 541 YNRASDYFCCKPEDYEGKNCNHLKHCRTPPCEVIDSCVTAMASNDTPEGVYISSNVCG 600
Qy 601 PHGCKKSQSGGKFTCDNKGFTGTGTYCHENINDCESNCRNGGTCIDGVNSYKICSDGWE 660
Db 601 PHGCKKSQSGGKFTCDNKGFTGTGTYCHENINDCESNCRNGGTCIDGVNSYKICSDGWE 660
Qy 661 GAYCETNINDCSQNPCHNGGTCDRLVNDFYCDCKNGWKGTCHSRDSQCDDEATCNNGGTC 720
Db 661 GAYCETNINDCSQNPCHNGGTCDRLVNDFYCDCKNGWKGTCHSRDSQCDDEATCNNGGTC 720
Qy 721 YDEGDAFKMCPGWEGETTCNIARNSSCLNPCHNGGTCTVNGESFTVCCKEGWGPICA 780
Db 721 YDEGDAFKMCPGWEGETTCNIARNSSCLNPCHNGGTCTVNGESFTVCCKEGWGPICA 780
Qy 781 QNTNDCSPHPCYNSGTCTVGDGNWYRCAPGAGPDCRININECQSSPCAFGATCVDEIN 840
Db 781 QNTNDCSPHPCYNSGTCTVGDGNWYRCAPGAGPDCRININECQSSPCAFGATCVDEIN 840
Qy 841 GYRCVCPGSHGAKQCVSRPCITMGSVIPDGAKWDDCNTCQCLNGRIACSKVWCGPR 900
Db 841 GYRCVCPGSHGAKQCVSRPCITMGSVIPDGAKWDDCNTCQCLNGRIACSKVWCGPR 900
Qy 901 PCLLKHGSHCPGQSCIPILDDQCFVHPTGTGVECRSSSLQPVKTKTSDSYQNCAN 960
Db 901 PCLLKHGSHCPGQSCIPILDDQCFVHPTGTGVECRSSSLQPVKTKTSDSYQNCAN 960
Qy 961 ITFTFNKEMMSPGLTTEHICSELRLNLIKNVSAEYSIYIACRSPSANNIEHVAISAED 1020
Db 961 ITFTFNKEMMSPGLTTEHICSELRLNLIKNVSAEYSIYIACRSPSANNIEHVAISAED 1020
Qy 1021 IRDGNPIKEITDKIIDLVSKRDNSSLIAAVAEVRVQRPLKRTD 1067
Db 1021 IRDGNPIKEITDKIIDLVSKRDNSSLIAAVAEVRVQRPLKRTD 1067
```

```
US-09-855-722-7
; Sequence 7, Application US/09855722
; Patent No. 6638741
; GENERAL INFORMATION:
; APPLICANT: Sakano, Seiji
; APPLICANT: Itoh, Akira
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8576
; CURRENT APPLICATION NUMBER: US/09/855,722
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/214,278
; PRIOR FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 7
; LENGTH: 1218
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-855-722-7
```

```
Query Match 99.7%; Score 6228; DB 2; Length 1218;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1065; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 MRSPTTRGRSRPLSLALLCALRAKVCASGQFELILSMQNVGELQNGCCGARN 60
Db 1 MRSPTTRGRSRPLSLALLCALRAKVCASGQFELILSMQNVGELQNGCCGARN 60
Qy 61 PGDRKCTRDECDTYFKVCLKEYQSRVTAGGPCSGSGSTPVIGNTFNKASRGNDRRI 120
Db 61 PGDRKCTRDESDTYFKVCLKEYQSRVTAGGPCSGSGSTPVIGNTFNKASRGNDRRI 120
Qy 121 VLPSPFAPRSYTLLEAVDSSNDTVQPDSEIEKASHGMINPSPQWTLKQNTGVVAHPE 180
Db 121 VLPSPFAPRSYTLLEAVDSSNDTVQPDSEIEKASHGMINPSPQWTLKQNTGVVAHPE 180
Qy 181 YQIRVTCDDYYFGCNKFCRPRDDFFGHYACDQNGKTCMEGWMGPECNRAICRQCSP 240
Db 181 YQIRVTCDDYYFGCNKFCRPRDDFFGHYACDQNGKTCMEGWMGPECNRAICRQCSP 240
Qy 241 KHGSKCLPGDCRCQYQWGLYCDKCI PHPGCVHGINPQCLCETNMGWGLCDKDLNYC 300
Db 241 KHGSKCLPGDCRCQYQWGLYCDKCI PHPGCVHGINPQCLCETNMGWGLCDKDLNYC 300
Qy 301 GTHQPCINGTCSNTGPDKYQCSCEGYSGPNCIEAHACLSDPCHNRGSKETSLSGFEC 360
Db 301 GTHQPCINGTCSNTGPDKYQCSCEGYSGPNCIEAHACLSDPCHNRGSKETSLSGFEC 360
Qy 361 ECSPGWTGPTCSNTIDDCSPNNSHGTCQDLVNGFKVCVCPPOWTGKTCOLDANECEAKP 420
Db 361 ECSPGWTGPTCSNTIDDCSPNNSHGTCQDLVNGFKVCVCPPOWTGKTCOLDANECEAKP 420
Qy 421 CVNAKSKNLIASYCDCLPGWQONCDININDCLGQONCDASCRDLVNGYRCICPPGYA 480
Db 421 CVNAKSKNLIASYCDCLPGWQONCDININDCLGQONCDASCRDLVNGYRCICPPGYA 480
Qy 481 GDHCERDIDECASNPCLNGHCHQNEINRFQCLCPTGFSGNLCOLDIDYCEPNPCQNGAQC 540
Db 481 GDHCERDIDECASNPCLNGHCHQNEINRFQCLCPTGFSGNLCOLDIDYCEPNPCQNGAQC 540
Qy 541 YNRASDYFCCKPEDYEGKNCNHLKHCRTPPCEVIDSCVTAMASNDTPEGVYISSNVCG 600
Db 541 YNRASDYFCCKPEDYEGKNCNHLKHCRTPPCEVIDSCVTAMASNDTPEGVYISSNVCG 600
Qy 601 PHGCKKSQSGGKFTCDNKGFTGTGTYCHENINDCESNCRNGGTCIDGVNSYKICSDGWE 660
Db 601 PHGCKKSQSGGKFTCDNKGFTGTGTYCHENINDCESNCRNGGTCIDGVNSYKICSDGWE 660
Qy 661 GAYCETNINDCSQNPCHNGGTCDRLVNDFYCDCKNGWKGTCHSRDSQCDDEATCNNGGTC 720
Db 661 GAYCETNINDCSQNPCHNGGTCDRLVNDFYCDCKNGWKGTCHSRDSQCDDEATCNNGGTC 720
Qy 721 YDEGDAFKMCPGWEGETTCNIARNSSCLNPCHNGGTCTVNGESFTVCCKEGWGPICA 780
```

```
Db 721 YDEGAFKCPGWEGTTCNIARNSSCLNPNCHNGTTCVNGESFTVCCKEWEGPICA 780
Qy 781 QNTNDCSPHPCYNSTGTCVGDGNWYRCECAFAGPDCRININECQSSPCAFGATCVDEN 840
Db 781 QNTNDCSPHPCYNSTGTCVGDGNWYRCECAFAGPDCRININECQSSPCAFGATCVDEN 840
Qy 841 GYRCVCPGHSAGAKQEVSRPCITMGSVIPDGAKWDDDCNTCOCLNGRIACSKWCGPR 900
Db 841 GYRCVCPGHSAGAKQEVSRPCITMGSVIPDGAKWDDDCNTCOCLNGRIACSKWCGPR 900
Qy 901 PCLLKHGSHSEPCSQSCIPILDDQCFVHPCTGVGECRSSLSLPVTKTCTSDSYQDNMCAN 960
Db 901 PCLLKHGSHSEPCSQSCIPILDDQCFVHPCTGVGECRSSLSLPVTKTCTSDSYQDNMCAN 960
Qy 961 ITFTFNKEMSPGLTTEHCSELNMLNKNVSAEYSIYIACEPSPSANNIEHVAISAED 1020
Db 961 ITFTFNKEMSPGLTTEHCSELNMLNKNVSAEYSIYIACEPSPSANNIEHVAISAED 1020
Qy 1021 IRDDGNPIKEITDKIIDLVSKRDNSSLIAAVAEVRVORRPLKNRTD 1067
Db 1021 IRDDGNPIKEITDKIIDLVSKRDNSSLIAAVAEVRVORRPLKNRTD 1067

RESULT 15
US-09-199-865-1
; Sequence 1, Application US/09199865
; Patent No. 6433138
; GENERAL INFORMATION:
; APPLICANT: Zimrin, Ann B.
; APPLICANT: Maciag, Thomas
; APPLICANT: Wong, Michael K.K.
; APPLICANT: Pepper, Michael S.
; APPLICANT: Montesano, Roberto
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC METHODS AND COMPOSITIONS
; TITLE OF INVENTION: BASED ON JAGGED/NOTCH PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: 0036-1U1
; CURRENT APPLICATION NUMBER: US/09/199,865
; CURRENT FILING DATE: 1998-11-25
; EARLIER APPLICATION NUMBER: 60/018,841
; EARLIER FILING DATE: 1996-05-31
; EARLIER APPLICATION NUMBER: PCT/US97/09407
; EARLIER FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-199-865-1

Query Match 98.8%; Score 6176; DB 2; Length 1208;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1054; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 11 GRPLSLLALLCALRAKVCASGQPELEILSMQNVNGBELQNGCCGARNPGRDKCTRDE 70
Db 1 GTSLSLLALLCALRAKVCASGQPELEILSMQNVNGBELQNGCCGARNPGRDKCTRDE 60
Qy 71 CDTVPKVCLEKQSVRTAGGPCSGSGSTPVIIGNTFNLKASRGNDRNRIVLPPSFAPWR 130
Db 61 CDTVPKVCLEKQSVRTAGGPCSGSGSTPVIIGNTFNLKASRGNDRNRIVLPPSFAPWR 120
Qy 131 SYTLIVEAWDSSNDTVQSDSIIEKASHSGMINPSRQWQTLKQNTGVAHFEYQIRVTCDDY 190
Db 121 SYTLIVEAWDSSNDTVQSDSIIEKASHSGMINPSRQWQTLKQNTGVAHFEYQIRVTCDDY 180
Qy 191 YYGFCNFKCPRDDFFGHYACDQNGKTCMEGWMGPCNRAICRQGSPKHGSKCLPGD 250
Db 181 YYGFCNFKCPRDDFFGHYACDQNGKTCMEGWMGPCNRAICRQGSPKHGSKCLPGD 240
Qy 251 CRCQYGMQGLYCDKCIPIHPCGVHGI CNEPWQCLCETNMGQGLCDKDLNYCGTHOPCLNGG 310
```

```
Db 241 CRCQYGMQGLYCDKCIPIHPCGVHGI CNEPWQCLCETNMGQGLCDKDLNYCGTHOPCLNGG 300
Qy 311 TCSNTGPDKYQCSCEPGYSGPNCEIAEHACLSDPCHNRGSKCKETSIGFECESPGWTGPT 370
Db 301 TCSNTGPDKYQCSCEPGYSGPNCEIAEHACLSDPCHNRGSKCKETSIGFECESPGWTGPT 360
Qy 371 CSTNIDDCSPNNCSHGCTCQDLVNGFKVCVCPPOWTGKTQOLDANECEAKPCVNAKSCNKL 430
Db 361 CSTNIDDCSPNNCSHGCTCQDLVNGFKVCVCPPOWTGKTQOLDANECEAKPCVNAKSCNKL 420
Qy 431 IASYYCDCLPFGMWGQNCIDINIDCLQOQONQDASCRDLVNGYRCICPPGVAGDHCERDIDE 490
Db 421 IASYYCDCLPFGMWGQNCIDINIDCLQOQONQDASCRDLVNGYRCICPPGVAGDHCERDIDE 480
Qy 491 CASNPCLNGHCHQNEINRFQCLCPTGFSGNLCOLDIDYCEPNPCONGAQCYNRSADYFCK 550
Db 481 CASNPCLNGHCHQNEINRFQCLCPTGFSGNLCOLDIDYCEPNPCONGAQCYNRSADYFCK 540
Qy 551 CPEDYEGKNCSHLKHCRITTPCEVIDSDCTVAMASNDTPCEGVYISSNVCGPHGCKKSQSG 610
Db 541 CPEDYEGKNCSHLKHCRITTPCEVIDSDCTVAMASNDTPCEGVYISSNVCGPHGCKKSQSG 600
Qy 611 GKFTCDNKGFTCTYCHENINDCESNPCRNGGTCIDGVNSYKICISDGWEGAYCETNIND 670
Db 601 GKFTCDNKGFTCTYCHENINDCESNPCRNGGTCIDGVNSYKICISDGWEGAYCETNIND 660
Qy 671 CSQNPCHNGGTCRDLVNDYFCDCKNGWKGTCHSRDSQCDCEATCNGGTCYDEGDAPFKCM 730
Db 661 CSQNPCHNGGTCRDLVNDYFCDCKNGWKGTCHSRDSQCDCEATCNGGTCYDEGDAPFKCM 720
Qy 731 CPGWEGTTCNIARNSSCLNPNCHNGGTCVNGESFTVCCKEWEGPICANTNDSCSPHP 790
Db 721 CPGWEGTTCNIARNSSCLNPNCHNGGTCVNGESFTVCCKEWEGPICANTNDSCSPHP 780
Qy 791 CYNSTGTCVGDGNWYRCECAFAGPDCRININECQSSPCAFGATCVDENGYRCVCPGPH 850
Db 781 CYNSTGTCVGDGNWYRCECAFAGPDCRININECQSSPCAFGATCVDENGYRCVCPGPH 840
Qy 851 SGAKQEVSRGPCITMGSVIPDGAKWDDDCNTCOCLNGRIACSKWCGPRPCLLKHGSHSE 910
Db 841 SGAKQEVSRGPCITMGSVIPDGAKWDDDCNTCOCLNGRIACSKWCGPRPCLLKHGSHSE 900
Qy 911 CPSSQSCIPILDDQCFVHPCTGVGECRSSLSLPVTKTCTSDSYQDNMCANITFTFNKEMM 970
Db 901 CPSSQSCIPILDDQCFVHPCTGVGECRSSLSLPVTKTCTSDSYQDNMCANITFTFNKEMM 960
Qy 971 SPGLTTEHCSELNMLNKNVSAEYSIYIACEPSPSANNIEHVAISAEDIRDDGNPIKE 1030
Db 961 SPGLTTEHCSELNMLNKNVSAEYSIYIACEPSPSANNIEHVAISAEDIRDDGNPIKE 1020
Qy 1031 ITDKIIDLVSKRDNSSLIAAVAEVRVORRPLKNRTD 1067
Db 1021 ITDKIIDLVSKRDNSSLIAAVAEVRVORRPLKNRTD 1057
```

Search completed: February 4, 2006, 09:37:26

Job time : 56 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 09:27:34 ; Search time 271 Seconds
(without alignments)

2777.855 Million cell updates/sec

Title: US-10-650-650-18

Perfect score: 6248

Sequence: 1 MSRPTRGRSGRPLSLALL.....LIAAENVRVQRPLKRTD 1067

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Uniprot 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6248	100.0	1218	1 JAG1_HUMAN	P78504 homo sapien
2	6239	99.9	1218	2 Q4KMR2_HUMAN	Q4KMR2 homo sapien
3	6089	97.5	1219	1 JAG1_RAT	Q63722 rattus norv
4	6086	97.4	1218	1 JAG1_MOUSE	Q9QX00 mus musculu
5	5469.5	87.5	1193	2 Q90819_CHICK	Q90819 gallus gall
6	5263	84.2	1214	2 Q90YD2_XENLA	Q90YD2 xenopus lae
7	5225	83.6	1051	2 Q5U4U1_XENLA	Q5U4U1 xenopus lae
8	4768.5	76.3	1213	1 JAG1B_BRARE	Q90Y54 brachydanio
9	4505.5	72.1	1242	1 JAG1A_BRARE	Q90Y57 brachydanio
10	4293.5	68.7	1364	2 Q4R003_TETNG	Q4R003 tetraodon n
11	3985	63.8	1212	2 Q42347_CHICK	Q42347 gallus gall
12	3795.5	60.7	1254	2 Q90Y56_BRARE	Q90Y56 brachydanio
13	3790.5	60.7	1254	2 Q5Y2K8_BRARE	Q5Y2K8 brachydanio
14	3789.5	60.7	1254	2 Q9YH22_BRARE	Q9YH22 brachydanio
15	3685	59.0	1238	1 JAG2_HUMAN	Q9Y219 homo sapien
16	3656	58.5	1247	1 JAG2_MOUSE	Q9QY65 mus musculu
17	3590.5	57.5	1216	2 Q90Y55_BRARE	Q90Y55 brachydanio
18	3585.5	57.4	1216	2 Q5Y2K7_BRARE	Q5Y2K7 brachydanio
19	3557.5	56.9	1202	1 JAG2_RAT	P97607 rattus norv
20	3174	50.8	1128	2 Q4SG68_TETNG	Q4SG68 tetraodon n
21	2923	46.8	1477	2 Q4H3A4_CIOIN	Q4H3A4 ciona intes
22	2335.5	37.4	1404	1 SERR_DROME	P18168 drosophila
23	2040.5	32.7	950	2 Q8MQN5_DROME	Q8MQN5 drosophila
24	1982.5	31.7	932	2 Q7Q6B7_ANOGA	Q7Q6B7 anopheles g
25	1585.5	25.4	2524	1 NOTCH_XENLA	P21783 xenopus lae
26	1584	25.4	2428	2 Q8I6X6_BOOMI	Q8I6X6 boophilus m
27	1582	25.3	2602	2 Q7PSV8_ANOGA	Q7PSV8 anopheles g
28	1580.5	25.3	2703	1 NOTCH_DROME	P07207 drosophila
29	1569.5	25.1	2528	2 Q8AXP0_CYNPY	Q8AXP0 cynops pyrr
30	1567	25.1	2471	1 NOTC2_HUMAN	Q04721 homo sapien
31	1567	25.1	2471	2 Q5VTD0_HUMAN	Q5VTD0 homo sapien

RESULT 1
JAG1_HUMAN
ID JAG1_HUMAN STANDARD; PRT; 1218 AA.
AC P78504; O14902; O15122; Q15816;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Jagged-1 precursor (Jagged1) (hJ1).
OS Homo sapiens (Human).
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RX MEDLINE=97422615; PubMed=9268641; DOI=10.1006/geno.1997.4820;
RA Oda T., Elkahlon A.G., Meltzer P.S., Chandrasekharappa S.C.;
RT "Identification and cloning of the human homolog (JAG1) of the rat
Jagged1 gene from the Alagille syndrome critical region at 20p12.";
RL Genomics 43:376-379(1997).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT AGS CVS-184.
RC TISSUE=Bone marrow;
RX MEDLINE=97351506; PubMed=9207788;
RA Li L., Krantz I.D., Deng Y., Genin A., Banta A.B., Collins C.C.,
Qi M., Trask B.J., Kuo W.L., Cochran J., Costa T., Pierpont M.E.M.,
Rand E.B., Piccoli D.A., Hood L., Spinner N.B.;
RT "Alagille syndrome is caused by mutations in human Jagged1, which
encodes a ligand for Notch1.";
RL Nat. Genet. 16:243-251(1997).
RN [3]
RP NUCLEOTIDE SEQUENCE [MRNA], AND FUNCTION.
RX MEDLINE=98122342; PubMed=9462510; DOI=10.1016/S1074-7613(00)80457-4;
RA Li L., Wilner L.A., Deng Y., Iwata M., Banta A.B., Graf L.,
Marcovina S., Friedman C., Trask B.J., Hood L., Torok-Storb B.;
RT "The human homolog of rat Jagged1 expressed by marrow stroma inhibits
differentiation of 32D cells through interaction with Notch1.";
RL Immunity 8:43-55(1998).
RN [4]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC TISSUE=Cervical carcinoma;
RX MEDLINE=99262417; PubMed=10329626; DOI=10.1093/emboj/18.10.2803;
RA Bash J., Zong W.-X., Banga S., Rivera A., Ballard D.W., Ron Y.,
Gelinias C.;
RT "Rel/NF-kappaB can trigger the Notch signaling pathway by inducing the
expression of Jagged1, a ligand for Notch receptors.";
RL EMBO J. 18:2803-2811(1999).
RN [5]
RP NUCLEOTIDE SEQUENCE [MRNA].
RX MEDLINE=99180765; PubMed=10079256;
RA Gray G.E., Mann R.S., Mitsiadis E., Henrique D., Carcangiu M.-L.,
Banks A., Leiman J., Ward D., Ieh-Horowitz D., Artavanis-Tsakonas S.;
RT "Human ligands of the Notch receptor.";

ALIGNMENTS

- Am. J. Pathol. 154:785-794(1999).
- [6]
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
MEDLINE=21638749; PubMed=11780052; DOI=10.1038/414865a;
Deloukas P., Matthews L.H., Ashurst J.L., Burton J., Balli F.,
Jones M., Scavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
Lehvaeslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
Marsh V.L., Martin S.L., McConachie L.J., McIlroy K., McMurray A.A.,
Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
Skane C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Williams S.A.,
Whitehead S.B., Whittaker P., Willey D.B., Williams L., Beck S.,
Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Rogers J.,
Rogers J.,
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
- [7]
NUCLEOTIDE SEQUENCE [MRNA] OF 14-1218.
TSSUP=Unilical vein endothelial cell;
MEDLINE=97115768; PubMed=8955070; DOI=10.1074/jbc.271.51.32499;
Zimin A.B., Pepper M.S., McMahon G.A., Nguyen F., Montesano R.,
Maciag T.,
RT "An antisense oligonucleotide to the notch ligand jagged enhances
RT fibroblast growth factor-induced angiogenesis in vitro.";
RL J. Biol. Chem. 271:32499-32502(1996).
- [8]
DISEASE.
MEDLINE=97351505; PubMed=9207787;
Oda T., Elkahoun A.G., Pike B.L., Okajima K., Krantz I.D., Genin A.,
Piccoli D.A., Meltzer P.S., Spinner N.B., Collins F.S.,
RA Chandrasekharappa S.C.;
RT "Mutations in the human Jagged1 gene are responsible for Alagille
RT syndrome.";
RL Nat. Genet. 16:235-242(1997).
- [9]
DEVELOPMENTAL STAGE.
MEDLINE=20436345; PubMed=10978356; DOI=10.1136/jmg.37.9.658;
Jones E.A., Clement-Jones M., Wilson D.I.;
RT "Jagged1 expression in human embryos: correlation with the Alagille
RT syndrome phenotype.";
RL J. Med. Genet. 37:663-668(2000).
- [10]
VARIANTS AGS CYS-184 AND HIS-184.
MEDLINE=98254456; PubMed=9585603;
Krantz I.D., Colliton R.P., Genin A., Rand E.B., Li L., Piccoli D.A.,
RA Spinner N.B.;
RT "Spectrum and frequency of jagged1 (JAG1) mutations in Alagille
RT syndrome patients and their families.";
RL Am. J. Hum. Genet. 62:1361-1369(1998).
- [11]
VARIANTS AGS HIS-79; THR-127; ARG-129; LEU-163; GLY-184; SER-187;
GLY-229; PHE-284; CYS-288; PHE-438; SER-731 AND ARG-740.
MEDLINE=99239888; PubMed=10220506;
RA Crosnier C., Driancourt C., Raynaud N., Dhorne-Pollet S., Pollet N.,
RA Bernard O., Hadchouel M., Meunier-Rotival M.;
RT "Mutations in JAGGED1 gene are predominantly sporadic in Alagille
RT syndrome.";
RL Gastroenterology 116:1141-1148(1999).
- [12]
VARIANTS AGS THR-152 AND LEU-184.
MEDLINE=20004539; PubMed=10533065;
DOI=10.1002/(SICI)1098-1004(199911)14:5<394::AID-HUMU5>3.0.CO;2-1;
RA Filia G., Uda M., Macis D., Frau F., Cripponi L., Balli F.,
RA Barbera C., Colombo C., Frediani T., Gatti R., Iorio R., Marazzi M.G.,
RA Marcellini M., Musumeci S., Nebbia G., Vajro P., Ruffa G., Zancan L.,
RA Cao A., Devirgilis S.;
RT "Jagged-1 mutation analysis in Italian Alagille syndrome patients.";
RL Hum. Mutat. 14:394-400(1999).
- [13]
VARIANTS AGS TYR-229 AND ARG-386.
MEDLINE=20514559; PubMed=11058898;
DOI=10.1002/1098-1004(200011)16:5<408::AID-HUMU5>3.0.CO;2-9;
RA Heritage M.L., MacMillan J.C., Colliton R.P., Genin A., Spinner N.B.,
RA Anderson G.J.;
RT "Jagged1 (JAG1) mutation detection in an Australian Alagille syndrome
RT population.";
RL Hum. Mutat. 16:408-416(2000).
- [14]
VARIANT TOP ASP-274.
MEDLINE=21067871; PubMed=11152664; DOI=10.1093/hmg/10.2.163;
RA Eldadah Z.A., Hamosh A., Biery N.J., Montgomery R.A., Duke M.,
RA Etkins R., Dietz H.C.;
RT "Familial Tetralogy of Fallot caused by mutation in the jagged1
RT gene.";
RL Hum. Mol. Genet. 10:163-169(2001).
- [15]
VARIANT AGS SER-37.
MEDLINE=21096916; PubMed=11157803; DOI=10.1093/hmg/10.4.405;
RA Morrisette J.D., Colliton R.P., Spinner N.B.;
RT "Defective intracellular transport and processing of JAG1 missense
RT mutations in Alagille syndrome.";
RL Hum. Mol. Genet. 10:405-413(2001).
- [16]
VARIANTS AGS PHE-220 AND ARG-753.
MEDLINE=20579880; PubMed=11139247;
DOI=10.1002/1098-1004(2001)17:1<72::AID-HUMU11>3.0.CO;2-U;
RA Crosnier C., Driancourt C., Raynaud N., Hadchouel M.,
RA Meunier-Rotival M.;
RT "Fifteen novel mutations in the JAGGED1 gene of patients with Alagille
RT syndrome.";
RL Hum. Mutat. 17:72-73(2001).
- [17]
FUNCTION: Ligand for multiple Notch receptors and involved in the
mediation of Notch signaling. May be involved in cell-fate
decisions during hematopoiesis. Seems to be involved in early and
late stages of mammalian cardiovascular development. Inhibits
myoblast differentiation (By similarity). Enhances fibroblast
growth factor-induced angiogenesis (in vitro).
SUBUNIT: Interacts with NOTCH1, NOTCH2 and NOTCH3 (By similarity).
SUBCELLULAR LOCATION: Type I membrane protein.
TISSUE SPECIFICITY: Widely expressed in adult and fetal tissues.
In cervix epithelium expressed in undifferentiated subcolumnar
reserve cells and squamous metaplasia. Expression is up-regulated
in cervical squamous cell carcinoma. Expressed in bone marrow cell
line HS-27a which supports the long-term maintenance of immature
progenitor cells.
DEVELOPMENTAL STAGE: Expressed in 32-52 days embryos in the distal
cardiac outflow tract and pulmonary artery, major arteries, portal
vein, optic vesicle, otocyst, branchial arches, metanephros,
pancreas, mesocardium, around the major bronchial branches, and in
the neural tube.
DISEASE: Defects in JAG1 are the cause of Alagille syndrome (AGS)
[MIM:118450]. AGS is an autosomal dominant developmental disorder
that affects structures in the liver, heart, skeleton, eye, kidney
and other organs.
DISEASE: Defects in JAG1 are a cause of tetralogy of Fallot (TOF)
[MIM:187500]. TOF is a congenital heart anomaly which consists of
pulmonary stenosis, ventricular septal defect, dextroposition of
the aorta (aorta is on the right side instead of the left) and
hypertrophy of the right ventricle. This condition results in a
blue baby at birth due to inadequate oxygenation. Surgical
correction is emergent.


```
QY 241 KHGSKCLPGBRCQYQWGLYCDKCIHPGCVHGIQNEPMQCLCETNWGGQLCKDLNYC 300
Db 241 KHGSKCLPGBRCQYQWGLYCDKCIHPGCVHGIQNEPMQCLCETNWGGQLCKDLNYC 300
QY 301 GTHQPCNLGGTCSNTGPDYKQCSPEGYSGPNCIEAHPACLSDPCHNRGSKETSGLGFEC 360
Db 301 GTHQPCNLGGTCSNTGPDYKQCSPEGYSGPNCIEAHPACLSDPCHNRGSKETSGLGFEC 360
QY 361 ECSPGWTPCTSNIDDCSPNCSHGTCQDLVNGFKVCYCPQWTKTQOLDANECEAKP 420
Db 361 ECSPGWTPCTSNIDDCSPNCSHGTCQDLVNGFKVCYCPQWTKTQOLDANECEAKP 420
QY 421 CVNAKSKNLIASYCDLPGWQNGCDIINDCLGQONDCASCRDLVNGYRCLCPGYA 480
Db 421 CVNAKSKNLIASYCDLPGWQNGCDIINDCLGQONDCASCRDLVNGYRCLCPGYA 480
QY 481 GDHCERDIDECASNPLCNGHGCQNEINRFQCLPTGFGSGLCQLDIDYCEPNFCQGAQC 540
Db 481 GDHCERDIDECASNPLCNGHGCQNEINRFQCLPTGFGSGLCQLDIDYCEPNFCQGAQC 540
QY 541 YNRASDYFCCKPEDYEGKNSHLKDHCRTPPCEVIDSCITVAMASNDTPGVRVYISSNVCG 600
Db 541 YNRASDYFCCKPEDYEGKNSHLKDHCRTPPCEVIDSCITVAMASNDTPGVRVYISSNVCG 600
QY 601 PHGCKKSQSGKFTCDCKNGKFTGTGTYCHENINDCESNRCNGGTCIDGWSYKICSDGWE 660
Db 601 PHGCKKSQSGKFTCDCKNGKFTGTGTYCHENINDCESNRCNGGTCIDGWSYKICSDGWE 660
QY 661 GAYCETNINDCSQNPCHNGGTCDRLVNDYFCDCKNGWKGTCHSRDSQCDATCNGGTC 720
Db 661 GAYCETNINDCSQNPCHNGGTCDRLVNDYFCDCKNGWKGTCHSRDSQCDATCNGGTC 720
QY 721 YDEGDAFKMCPGWEGTTCNIARNSSCLNPNCHNGTCTVNGESFTVCYCKEGWGPICA 780
Db 721 YDEGDAFKMCPGWEGTTCNIARNSSCLNPNCHNGTCTVNGESFTVCYCKEGWGPICA 780
QY 781 QNTNDCSPHPCYNSTGTCVGDWNYRCAFGAPGDCRINIINECQSPCAFGATCVDEIN 840
Db 781 QNTNDCSPHPCYNSTGTCVGDWNYRCAFGAPGDCRINIINECQSPCAFGATCVDEIN 840
QY 841 GYRCVCPFGHSGAKQEVSRPCITMTGSVIPDGAkWDDDCNTCQCLNGRIACSKVMCGPR 900
Db 841 GYRCVCPFGHSGAKQEVSRPCITMTGSVIRDGAkWDDDCNTCQCLNGRIACSKVMCGPR 900
QY 901 PCLLKHGSECPQSQCIPILDDQCFVHPTGVCESRSSLOPVKTKTSDSYQDNKAN 960
Db 901 PCLLKHGSECPQSQCIPILDDQCFVHPTGVCESRSSLOPVKTKTSDSYQDNKAN 960
QY 961 ITFTFNKEMWSPGLTTEHICSELRLNLIKNVSAEYIYIACEPSPANNEIHVAISAED 1020
Db 961 ITFTFNKEMWSPGLTTEHICSELRLNLIKNVSAEYIYIACEPSPANNEIHVAISAED 1020
QY 1021 IRDGNPIKEITDKIIDLVSQRDNSSLIAAVAEVRVQRPLKNRDT 1067
Db 1021 IRDGNPIKEITDKIIDLVSQRDNSSLIAAVAEVRVQRPLKNRDT 1067

RESULT 3
JAG1 RAT
ID JAG1 RAT STANDARD; PRT; 1219 AA.
AC Q63722; P70640;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Jagged-1 precursor (Jagged1).
GN Name=Jagi;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
```

```
RC TISSUE=Sciatic nerve;
RX MEDLINE=95211842; PubMed=7697721; DOI=10.1016/0092-8674(95)90294-5;
RA Lindell C.B., Shawer C.J., Boulter J., Weinmaster G.;
RT "Jagged: a mammalian ligand that activates Notch1.";
RL Cell 80:909-917(1995).
CC -1- FUNCTION: Ligand for multiple Notch receptors and involved in the
CC mediation of Notch signaling. May be involved in cell-fate
CC decisions during hematopoiesis. Enhances fibroblast growth factor-
CC induced angiogenesis (in vitro). Seems to be involved in early and
CC late stages of mammalian cardiovascular development. Inhibits
CC myoblast differentiation. May regulate fibroblast growth factor-
CC induced angiogenesis.
CC -1- SUBUNIT: Interacts with NOTCH1, NOTCH2 and NOTCH3 (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Widely expressed in a variety of tissues.
CC -1- DEVELOPMENTAL STAGE: Expression is seen in E11.5-E14.5 embryos in
CC four distinct regions of the ventricular zone in the developing
CC spinal cord.
CC -1- SIMILARITY: Contains 1 DSL domain.
CC -1- SIMILARITY: Contains 15 EGF-like domains.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; L38483; AAB06509.1; -, mRNA.
DR HSSP; P00740; 1EDM.
DR Ensembl; ENSRNOG00000007443; Rattus norvegicus.
DR RGD; 2937; Jag1.
DR GO; GO:0005576; C:extracellular region; ISS.
DR GO; GO:0005887; C:integral to plasma membrane; ISS.
DR GO; GO:0008083; F:growth factor activity; ISS.
DR GO; GO:0005112; F:Notch binding; ISS.
DR GO; GO:0005198; F:structural molecule activity; ISS.
DR GO; GO:0001525; P:angiogenesis; ISS.
DR GO; GO:0001709; P:cell fate determination; NAS.
DR GO; GO:0045446; P:endothelial cell differentiation; ISS.
DR GO; GO:0030097; P:hemopoiesis; ISS.
DR GO; GO:0030216; P:keratinocyte differentiation; ISS.
DR GO; GO:0045445; P:myoblast differentiation; ISS.
DR GO; GO:0007399; P:neurogenesis; ISS.
DR GO; GO:0007219; P:Notch signaling pathway; ISS.
DR GO; GO:0042127; P:regulation of cell proliferation; ISS.
DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR001774; DSL.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR011651; MNFL.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF01414; DSL; 1.
DR Pfam; PF00008; EGF; 11.
DR Pfam; PF07645; EGF_CA; 1.
DR Pfam; PF07657; MNFL; 1.
DR PRINTS; PR00010; EGFLOOD.
DR SMART; SM00051; DSL; 1.
DR SMART; SM00179; EGF_CA; 10.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 10.
DR PROSITE; PS01051; DSL; 1.
DR PROSITE; PS00022; EGF_1; 16.
DR PROSITE; PS01186; EGF_2; 12.
DR PROSITE; PS00026; EGF_3; 16.
DR PROSITE; PS01187; EGF_CA; 8.
KW Calcium; Developmental protein; EGF-like domain; Glycoprotein;
KW Notch signaling pathway; Repeat; Signal; Transmembrane.
FT SIGNAL 1 33 Potential.
FT CHAIN 34 1219 Jagged-1.
FT TOPO_DOM 34 1067 Extracellular (Potential).
FT TRANSMEM 1068 1093 Potential.
```

FT	TOPO_DOM	1094	1219	Cytoplasmic (Potential).	QY	1	MSPRTRGRSRLPLSLALLCALRAKVCASGOFELEILSMQNVANGELQNGNCCGARN	60
FT	DOMAIN	185	229	DSL.	QY	1		
FT	DOMAIN	230	266	EGF-like 1; atypical.	Db	1		
FT	DOMAIN	296	334	EGF-like 2.	Db	1		
FT	DOMAIN	336	372	EGF-like 3.	QY	61	PGDRKTRDECDTYFKVCLKEVQSRVTAGGPCSFGSGSTPVIGNTFNLKASRGNDNRRI	120
FT	DOMAIN	374	410	EGF-like 4; calcium-binding (Potential).	QY	61		
FT	DOMAIN	412	448	EGF-like 5; calcium-binding (Potential).	QY	61	PGDRKTRDECDTYFKVCLKEVQSRVTAGGPCSFGSGSTPVIGNTFNLKASRGNDNRRI	120
FT	DOMAIN	450	485	EGF-like 6; calcium-binding (Potential).	QY	121	VLPFSFAWPRSYTLVLEAWDSSNDTTPQDSIIIEKASHGMINPSRQWTLKQNTGVAAHFE	180
FT	DOMAIN	487	523	EGF-like 7; calcium-binding (Potential).	QY	121		
FT	DOMAIN	525	561	EGF-like 8.	QY	121	VLPFSFAWPRSYTLVLEAWDSSNDTTPQDSIIIEKASHGMINPSRQWTLKQNTGVAAHFE	180
FT	DOMAIN	574	627	EGF-like 9.	QY	121		
FT	DOMAIN	629	665	EGF-like 10; calcium-binding (Potential).	QY	181	YQIRVTCDDYYGFGCNKFCRPRDDFFGHYACDQNGNKTCEMGWMPGECNRAICROGCSP	240
FT	DOMAIN	667	703	EGF-like 11; calcium-binding (Potential).	QY	181		
FT	DOMAIN	705	741	EGF-like 12.	QY	181	YQIRVTCDDYYGFGCNKFCRPRDDFFGHYACDQNGNKTCEMGWMPGECNRAICROGCSP	240
FT	DOMAIN	744	780	EGF-like 13.	QY	241	KHGSKLPGDCRCQYGMQGLYCDKCIHPHGCYVHGIENPWCICETNMGQOLCDKLDNYC	300
FT	DOMAIN	782	818	EGF-like 14; calcium-binding (Potential).	QY	241		
FT	DOMAIN	820	856	EGF-like 15; calcium-binding (Potential).	QY	241	KHGSKLPGDCRCQYGMQGLYCDKCIHPHGCYVHGTCEMGWMPGECNRAICROGCSP	300
FT	CARBOHYD	143	143	N-linked (GlnAc. . .) (Potential).	QY	301	GTQPCPLNGGTCNTGPDYKQSCPEGYSGPNCIEAIAHACLSDPCHNRSGCKETSIGFEC	360
FT	CARBOHYD	217	217	N-linked (GlnAc. . .) (Potential).	QY	301		
FT	CARBOHYD	382	382	N-linked (GlnAc. . .) (Potential).	QY	301	GTQPCPLNGGTCNTGPDYKQSCPEGYSGPNCIEAIAHACLSDPCHNRSGCKETSIGFEC	360
FT	CARBOHYD	559	559	N-linked (GlnAc. . .) (Potential).	QY	301		
FT	CARBOHYD	745	745	N-linked (GlnAc. . .) (Potential).	QY	301	GTQPCPLNGGTCNTGPDYKQSCPEGYSGPNCIEAIAHACLSDPCHNRSGCKETSIGFEC	360
FT	CARBOHYD	960	960	N-linked (GlnAc. . .) (Potential).	QY	361	ECSPGWTGPTCSNIDDCSPNNCSHGTCQDLVNGFKVCVPPQWTGKTCOLDANECEAKP	420
FT	CARBOHYD	991	991	N-linked (GlnAc. . .) (Potential).	QY	361		
FT	CARBOHYD	1045	1045	N-linked (GlnAc. . .) (Potential).	QY	361	ECSPGWTGPTCSNIDDCSPNNCSHGTCQDLVNGFKVCVPPQWTGKTCOLDANECEAKP	420
FT	CARBOHYD	1064	1064	N-linked (GlnAc. . .) (Potential).	QY	421	CYNAKSKNLIASYCDCLPGWNGQNDININDCLGQCCNDASCRDLVNGYRICIPGYA	480
FT	DISULFID	300	312	By similarity.	QY	421		
FT	DISULFID	306	322	By similarity.	QY	421	CYNAKSKNLIASYCDCLPGWNGQNDININDCLGQCCNDASCRDLVNGYRICIPGYA	480
FT	DISULFID	324	333	By similarity.	QY	481	GDHCERDIDECASNPCPLNGHGHQNEINRFQCLCPTGFSGNLCOLDIDYCEPNPCQNGAQC	540
FT	DISULFID	340	351	By similarity.	QY	481		
FT	DISULFID	345	360	By similarity.	QY	481	GDHCERDIDECASNPCPLNGHGHQNEINRFQCLCPTGFSGNLCOLDIDYCEPNPCQNGAQC	540
FT	DISULFID	362	371	By similarity.	QY	541	YNRASDYCKCPEDYEGKNCCHLKHCRTPPCVIDSCTVAMASNDTPEGVRYISSNVCG	600
FT	DISULFID	378	389	By similarity.	QY	541		
FT	DISULFID	383	398	By similarity.	QY	601	PHGCKSKSGGKFTCDCKNGFTGTGCHENINDCESNPCRNGGTCIDGVNSKYKICSDGWE	660
FT	DISULFID	400	409	By similarity.	QY	601		
FT	DISULFID	416	427	By similarity.	QY	601	PHGCKSKSGGKFTCDCKNGFTGTGCHENINDCESNPCRNGGTCIDGVNSKYKICSDGWE	660
FT	DISULFID	421	436	By similarity.	QY	661	GAYCETNINDCSQNPCHNGGTCRDVLVNDYFDCCKNGWKGTCHSRDSQCDCEATCNGGTC	720
FT	DISULFID	438	447	By similarity.	QY	661		
FT	DISULFID	454	464	By similarity.	QY	721	YDEGDAFKMCPGWEGETTCNIAARNSSCLPNPCHNGGTCVNGDSFTVCCKEGWGPIC	780
FT	DISULFID	458	473	By similarity.	QY	721		
FT	DISULFID	475	484	By similarity.	QY	781	QNTNDCSPHPCVNSGTCVVDGDNWYRCCECAPGAGPDCRININECQSSPCAFGATCVDEN	840
FT	DISULFID	491	502	By similarity.	QY	841	GYRCVCPHSGAKCOEVSGRPCITMGSVIPDGAKWDCCNTCCCLNGRITACSKVWCGPR	900
FT	DISULFID	511	511	By similarity.	QY	841		
FT	DISULFID	516	522	By similarity.	QY	841	GYRCVCPHSGAKCOEVSGRPCITMGSVIPDGAKWDCCNTCCCLNGRITACSKVWCGPR	900
FT	DISULFID	529	540	By similarity.	QY	901	PCLLKHGHECPNGQSCIPVLDQCFVHPCTGVGECRSSSLQPVKTKTSDSYQDNCAN	960
FT	DISULFID	534	549	By similarity.	QY	901		
FT	DISULFID	551	560	By similarity.	QY	961	ITFTFNKEMWSPGLTTEHICSELRLNLIKNVSAEYSIYIACPPSPSANNEIHVAISAED	1020
FT	DISULFID	578	605	By similarity.	QY	961	ITFTFNKEMWSPGLTTEHICSELRLNLIKNVSAEYSIYIACPPSPSANNEIHVAISAED	1020
FT	DISULFID	599	615	By similarity.	QY	1021	IRDDGNPIKEITDKIIDLVSKRDNSSLIATAAEVVRQRRPLKNRTD	1067
FT	DISULFID	615	626	By similarity.	QY	1021	IRDDGNPIKEITDKIIDLVSKRDNSSLIATAAEVVRQRRPLKNRTD	1067
FT	DISULFID	617	644	By similarity.	QY	1021	IRDDGNPIKEITDKIIDLVSKRDNSSLIATAAEVVRQRRPLKNRTD	1067
FT	DISULFID	633	644	By similarity.	QY	1021	IRDDGNPIKEITDKIIDLVSKRDNSSLIATAAEVVRQRRPLKNRTD	1067
FT	DISULFID	638	653	By similarity.	QY	1021	IRDDGNPIKEITDKIIDLVSKRDNSSLIATAAEVVRQRRPLKNRTD	1067
FT	DISULFID	655	664	By similarity.	QY	1021	IRDDGNPIKEITDKIIDLVSKRDNSSLIATAAEVVRQRRPLKNRTD	1067
FT	DISULFID	671	682	By similarity.	QY	1021	IRDDGNPIKEITDKIIDLVSKRDNSSLIATAAEVVRQRRPLKNRTD	1067
FT	DISULFID	676	691	By similarity.	QY	1021	IRDDGNPIKEITDKIIDLVSKRDNSSLIATAAEVVRQRRPLKNRTD	1067
FT	DISULFID	693	702	By similarity.	QY	1021	IRDDGNPIKEITDKIIDLVSKRDNSSLIATAAEVVRQRRPLKNRTD	1067
FT	DISULFID	709	720	By similarity.	QY	1021	IRDDGNPIKEITDKIIDLVSKRDNSSLIATAAEVVRQRRPLKNRTD	1067
FT	DISULFID	714	729	By similarity.	QY	1021	IRDDGNPIKEITDKIIDLVSKRDNSSLIATAAEVVRQRRPLKNRTD	1067
FT	DISULFID	731	740	By similarity.	QY	1021	IRDDGNPIKEITDKIIDLVSKRDNSSLIATAAEVVRQRRPLKNRTD	1067
FT	DISULFID	748	759	By similarity.	QY	1021	IRDDGNPIKEITDKIIDLVSKRDNSSLIATAAEVVRQRRPLKNRTD	1067
FT	DISULFID	753	768	By similarity.	QY	1021	IRDDGNPIKEITDKIIDLVSKRDNSSLIATAAEVVRQRRPLKNRTD	1067
FT	DISULFID	770	779	By similarity.	QY	1021	IRDDGNPIKEITDKIIDLVSKRDNSSLIATAAEVVRQRRPLKNRTD	1067
FT	DISULFID	786	797	By similarity.	QY	1021	IRDDGNPIKEITDKIIDLVSKRDNSSLIATAAEVVRQRRPLKNRTD	1067
FT	DISULFID	791	806	By similarity.	QY	1021	IRDDGNPIKEITDKIIDLVSKRDNSSLIATAAEVVRQRRPLKNRTD	1067
FT	DISULFID	808	817	By similarity.	QY	1021	IRDDGNPIKEITDKIIDLVSKRDNSSLIATAAEVVRQRRPLKNRTD	1067
FT	DISULFID	824	835	By similarity.	QY	1021	IRDDGNPIKEITDKIIDLVSKRDNSSLIATAAEVVRQRRPLKNRTD	1067
FT	DISULFID	829	844	By similarity.	QY	1021	IRDDGNPIKEITDKIIDLVSKRDNSSLIATAAEVVRQRRPLKNRTD	1067
FT	DISULFID	846	855	By similarity.	QY	1021	IRDDGNPIKEITDKIIDLVSKRDNSSLIATAAEVVRQRRPLKNRTD	1067
SQ	SEQUENCE	1219	AA; 134326 MW; 65D4CFC238A0E204	CRC64;	QY	1021	IRDDGNPIKEITDKIIDLVSKRDNSSLIATAAEVVRQRRPLKNRTD	1067

Query Match 97.5%; Score 6089; DB 1; Length 1219;
Best Local Similarity 96.8%; Pred. No. 0;
Matches 1033; Conservative 14; Mismatches 20; Indels 0; Gaps 0;

RESULT 4
JAG1_MOUSE STANDARD; PRT; 1218 AA.
AC Q9QXX0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Jagged-1 precursor (Jagged1).
GN Name=Jag1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN NUCLEOTIDE SEQUENCE, AND RECEPTOR INTERACTION.
RC STRAIN=Swiss Webster / NIH;
RX MEDLINE=20020271; PubMed=10551863; DOI=10.1074/jbc.274.46.32961;
RA Shimizu K., Chiba S., Kumano K., Hosoya N., Takahashi T., Kanda Y.,
RA Hamada Y., Iizaki Y., Hirai H.;
RT "Mouse Jagged1 physically interacts with Notch2 and other Notch
RT receptors: assessment by quantitative methods.";
RL J. Biol. Chem. 274:32961-32969(1999).
RN [2]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J.J., Helton E., Kettman M., Madan A.C., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RN TISSUE SPECIFICITY.
RX MEDLINE=20025753; PubMed=10556292; DOI=10.1093/hmg/8.13.2443;
RA Loomes K.M., Underkoffler L.A., Morabito J., Gottlieb S.,
RA Piccoli D.A., Spinner N.B., Baldwin H.S., Oakley R.J.;
RT "The expression of Jagged1 in the developing mammalian heart
RT correlates with cardiovascular disease in Alagille syndrome.";
RL Hum. Mol. Genet. 8:2443-2449(1999).
CC -1- FUNCTION: Ligand for multiple Notch receptors and involved in the
CC mediation of Notch signaling. May be involved in cell-fate
CC decisions during hematopoiesis. Seems to be involved in early and
CC late stages of mammalian cardiovascular development. Inhibits
CC myoblast differentiation (By similarity). May regulate fibroblast
CC growth factor-induced angiogenesis.
CC -1- SUBUNIT: Interacts with NOTCH1, NOTCH2 and NOTCH3.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Widely expressed in many tissues, with highest
CC expression in brain, heart, muscle and thymus.
CC -1- DEVELOPMENTAL STAGE: At 8.75-9.75 dpc expression was detected in
CC structures that include those destined to contribute to the
CC cardiovascular system of the adult heart. Expression was also
CC detected in the mesencephalon and rhombencephalon.
CC -1- DOMAIN: The DSL domain is indispensable and sufficient for binding
CC to NOTCH2.
CC -1- SIMILARITY: Contains 1 DSL domain.

CC -1- SIMILARITY: Contains 15 EGF-like domains.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AF171092; AAF15505.1; -; mRNA.
DR EMBL; BC058675; AAH58675.1; -; mRNA.
DR HSSP; P00740; IEDM.
DR Ensemble; ENSMUSG00000027276; Mus musculus.
DR MGI; MGI:1095416; Jag1.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR GO; GO:0005509; F:calcium ion binding; NAS.
DR GO; GO:0008083; F:growth factor activity; ISS.
DR GO; GO:0005112; F:Notch binding; IPI.
DR GO; GO:0005198; F:structural molecule activity; ISS.
DR GO; GO:001525; P:angiogenesis; ISS.
DR GO; GO:001709; P:cell fate determination; ISS.
DR GO; GO:0045446; P:endothelial cell differentiation; ISS.
DR GO; GO:0030097; P:hemoiesis; ISS.
DR GO; GO:0020216; P:keratinocyte differentiation; ISS.
DR GO; GO:0002011; P:morphogenesis of an epithelial sheet; IMP.
DR GO; GO:0045445; P:myoblast differentiation; ISS.
DR GO; GO:0045596; P:negative regulation of cell differentiation; IMP.
DR GO; GO:0007399; P:neurogenesis; ISS.
DR GO; GO:0007219; P:Notch signaling pathway; ISS.
DR GO; GO:0045639; P:positive regulation of myeloid cell differ. .; IDA.
DR GO; GO:0045747; P:positive regulation of Notch signaling pathway; IDA.
DR GO; GO:0042127; P:regulation of cell proliferation; ISS.
DR InterPro; IPR000152; Asx_hydroxyl_s.
DR InterPro; IPR001774; DSL.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR011651; MNML.
DR Pfam; PF01414; DSL; 1.
DR Pfam; PF00008; EGF; 11.
DR Pfam; PF07645; EGF_CA; 1.
DR Pfam; PF07657; MNML; 1.
DR PRINTS; PR00010; EGFBL00D.
DR PROSITE; PS00010; ASX_HYDROXYL; 10.
DR PROSITE; PS00022; EGF_1; 16.
DR PROSITE; PS01186; EGF_2; 12.
DR PROSITE; PS50026; EGF_3; 15.
DR PROSITE; PS01187; EGF_CA; 8.
DR Calcium; Developmental protein; EGF-like domain; Glycoprotein;
DR Notch signaling pathway; Repeat; Signal; Transmembrane.
FT SIGNAL 1 33 Potential.
FT CHAIN 34 1218 Jagged-1.
FT TOPO_DOM 34 1067 Extracellular (Potential).
FT TRANSMEM 1068 1093 Potential.
FT TOPO_DOM 1094 1218 Cytoplasmic (Potential).
FT DOMAIN 185 229 DSL.
FT DOMAIN 230 266 EGF-like 1; atypical.
FT DOMAIN 296 334 EGF-like 2.
FT DOMAIN 336 372 EGF-like 3.
FT DOMAIN 374 410 EGF-like 4; calcium-binding (Potential).
FT DOMAIN 412 448 EGF-like 5; calcium-binding (Potential).
FT DOMAIN 450 485 EGF-like 6; calcium-binding (Potential).
FT DOMAIN 487 523 EGF-like 7; calcium-binding (Potential).
FT DOMAIN 525 561 EGF-like 8.
FT DOMAIN 574 627 EGF-like 9.
FT DOMAIN 629 665 EGF-like 10; calcium-binding (Potential).
FT DOMAIN 667 703 EGF-like 11; calcium-binding (Potential).
FT DOMAIN 705 741 EGF-like 12.
FT DOMAIN 744 780 EGF-like 13.
FT DOMAIN 782 818 EGF-like 14; calcium-binding (Potential).
FT DOMAIN 820 856 EGF-like 15; calcium-binding (Potential).
CC

FT	CARBOHYD	143	143	N-linked (GlcNAc. . .)	(Potential).
FT	CARBOHYD	217	217	N-linked (GlcNAc. . .)	(Potential).
FT	CARBOHYD	382	382	N-linked (GlcNAc. . .)	(Potential).
FT	CARBOHYD	559	559	N-linked (GlcNAc. . .)	(Potential).
FT	CARBOHYD	745	745	N-linked (GlcNAc. . .)	(Potential).
FT	CARBOHYD	960	960	N-linked (GlcNAc. . .)	(Potential).
FT	CARBOHYD	991	991	N-linked (GlcNAc. . .)	(Potential).
FT	CARBOHYD	1045	1045	N-linked (GlcNAc. . .)	(Potential).
FT	CARBOHYD	1064	1064	N-linked (GlcNAc. . .)	(Potential).
FT	DISULFID	300	312	By similarity.	
FT	DISULFID	306	322	By similarity.	
FT	DISULFID	324	333	By similarity.	
FT	DISULFID	340	351	By similarity.	
FT	DISULFID	345	360	By similarity.	
FT	DISULFID	362	371	By similarity.	
FT	DISULFID	378	389	By similarity.	
FT	DISULFID	383	398	By similarity.	
FT	DISULFID	400	409	By similarity.	
FT	DISULFID	416	427	By similarity.	
FT	DISULFID	421	436	By similarity.	
FT	DISULFID	438	447	By similarity.	
FT	DISULFID	454	464	By similarity.	
FT	DISULFID	458	473	By similarity.	
FT	DISULFID	475	484	By similarity.	
FT	DISULFID	491	502	By similarity.	
FT	DISULFID	496	511	By similarity.	
FT	DISULFID	513	522	By similarity.	
FT	DISULFID	529	540	By similarity.	
FT	DISULFID	534	549	By similarity.	
FT	DISULFID	551	560	By similarity.	
FT	DISULFID	578	605	By similarity.	
FT	DISULFID	599	615	By similarity.	
FT	DISULFID	617	626	By similarity.	
FT	DISULFID	633	644	By similarity.	
FT	DISULFID	638	653	By similarity.	
FT	DISULFID	655	664	By similarity.	
FT	DISULFID	671	682	By similarity.	
FT	DISULFID	676	691	By similarity.	
FT	DISULFID	693	702	By similarity.	
FT	DISULFID	709	720	By similarity.	
FT	DISULFID	714	729	By similarity.	
FT	DISULFID	731	740	By similarity.	
FT	DISULFID	748	759	By similarity.	
FT	DISULFID	753	768	By similarity.	
FT	DISULFID	770	779	By similarity.	
FT	DISULFID	786	797	By similarity.	
FT	DISULFID	791	806	By similarity.	
FT	DISULFID	808	817	By similarity.	
FT	DISULFID	824	835	By similarity.	
FT	DISULFID	829	844	By similarity.	
FT	DISULFID	846	855	By similarity.	
SQ	SEQUENCE	1218 AA;	134164 MW; 77739F828BB793C CRC64;		
Query Match 97.48; Score 6086; DB 1; Length 1218;					
Best Local Similarity 96.84; Pred. No. 0;					
Matches 1033; Conservative 15; Mismatches 19; Indels 0; Gaps 0;					
Qy	1	MRSPTGRGRPLSLALLCALRAKVCASGQFEILSMQNVANGELQNGCCGARN	60		
Db	1	MRSPTGRGRPLSLALLCALRAKVCASGQFEILSMQNVANGELQNGCCGARN	60		
Qy	61	PGDRKCTRDECDTYFKVCLKEYQSRVTAGPCFSGGSTPVIQGNFTNLKASRGNDRNI	120		
Db	61	PGDRKCTRDECDTYFKVCLKEYQSRVTAGPCFSGGSTPVIQGNFTNLKASRGNDRNI	120		
Qy	121	VLPSFAWPRSYLLVEAWDSSNDTQPDSEIIEKASHGMINPSROWTLKQNTGVAHPE	180		
Db	121	VLPSFAWPRSYLLVEAWDSSNDTQPDSEIIEKASHGMINPSROWTLKQNTGVAHPE	180		
Qy	181	YQIRVTCDHYGFCNKFRCRDRDDPFGHYACDQNGKTCMEGWMGPECNRAICROGCSP	240		
Db	181	YQIRVTCDHYGFCNKFRCRDRDDPFGHYACDQNGKTCMEGWMGPECNRAICROGCSP	240		

Qy	241	XHGSCKLPBDCRCQYQWGLYCDKCI	PHPGCVHGI	CNEPQCICETNWGQGLCDKOLNYC	300
Db	241	XHGSCKLPBDCRCQYQWGLYCDKCI	PHPGCVHGI	CNEPQCICETNWGQGLCDKOLNYC	300
Qy	301	GTHQPCNLNGTCSNTGPKYQCSCEPGYSGPNCIEA	HACLSDPCHNRGSKETSLSGPEC	360	
Db	301	GTHQPCNLNGTCSNTGPKYQCSCEPGYSGPNCIEA	HACLSDPCHNRGSKETSLSGPEC	360	
Qy	361	ECSPGWTGPTCTNIDDCSPNNCSHGTCQDLVNGFKVCV	PPQWMTGKTCOLDANECEAKP	420	
Db	361	ECSPGWTGPTCTNIDDCSPNNCSHGTCQDLVNGFKVCV	PPQWMTGKTCOLDANECEAKP	420	
Qy	421	CYNAKSKNLIASYCDCLPGWNGQCDININCLGQCDNDAS	CRDLVNGYRCICPPGYA	480	
Db	421	CYNAKSKNLIASYCDCLPGWNGQCDININCLGQCDNDAS	CRDLVNGYRCICPPGYA	480	
Qy	481	GDHCERDIDECASNPLNGHGHQNEINRFOLCPTGFSGNL	COLDIDYCEPNPCQNGAQC	540	
Db	481	GDHCERDIDECASNPLNGHGHQNEINRFOLCPTGFSGNL	COLDIDYCEPNPCQNGAQC	540	
Qy	541	YNRASDYFCKCPEDYEGKNCSHLKHCRTRTPCEVIDS	CTVMAASNDTPEGVRYISSNVCG	600	
Db	541	YNRASDYFCKCPEDYEGKNCSHLKHCRTRTPCEVIDS	CTVMAASNDTPEGVRYISSNVCG	600	
Qy	601	PHGCKSQSGGKFTCDCKNGFTGTYCHENINDCESNP	CRNGGTCIDGWSYKICISDGWE	660	
Db	601	PHGCKSQSGGKFTCDCKNGFTGTYCHENINDCESNP	CRNGGTCIDGWSYKICISDGWE	660	
Qy	661	GAYCETNINDSONPCHNGTCDRLVNDFYCDCKNGWK	GKTKCHSRDSQCDDEATCNGGTC	720	
Db	661	GAYCETNINDSONPCHNGTCDRLVNDFYCDCKNGWK	GKTKCHSRDSQCDDEATCNGGTC	720	
Qy	721	YDEGDAFKMCPGGWEGTTCNIARNSSCLPNPCHNGT	CTVYNGESFTVCCKEGWEGPICA	780	
Db	721	YDEGDAFKMCPGGWEGTTCNIARNSSCLPNPCHNGT	CTVYNGESFTVCCKEGWEGPICA	780	
Qy	781	QNTNDCSHPHCYNSTGTCVGDNDWYRCAPAGP	DCRININECQSSPCAFCATCVDEN	840	
Db	781	QNTNDCSHPHCYNSTGTCVGDNDWYRCAPAGP	DCRININECQSSPCAFCATCVDEN	840	
Qy	841	GYRCVCPGHSGAKCOEVSGRPCITWGSVTPDCAKWD	DDCNTCQCLNGRACSKVWCGR	900	
Db	841	GYRCVCPGHSGAKCOEVSGRPCITWGSVTPDCAKWD	DDCNTCQCLNGRACSKVWCGR	900	
Qy	901	PCLLHKGHSECPSSQSCIPLDDQCFVHPCTGVGECR	SSLSQPVTKTCTSDSYQDNCA	960	
Db	901	PCLLHKGHSECPSSQSCIPLDDQCFVHPCTGVGECR	SSLSQPVTKTCTSDSYQDNCA	960	
Qy	961	ITFTFNKEMMSPGLTTEHICSELURNILKNVSAEYS	IYIACPPSPSANNIEHVAISAED	1020	
Db	961	ITFTFNKEMMSPGLTTEHICSELURNILKNVSAEYS	IYIACPPSPSANNIEHVAISAED	1020	
Qy	1021	IRDGNGPIKEITDKIIDLVSKRDNSSLTAAV	AEVVRQREPLKNRTD	1067	
Db	1021	IRDGNGPIKEITDKIIDLVSKRDNSSLTAAV	AEVVRQREPLKNRTD	1067	
RESULT 5					
Q90819_CHICK					
ID	Q90819	CHICK PRELIMINARY;	PRT; 1193 AA.		
AC	Q90819				
DT	01-NOV-1996	(TrEMBLrel. 01, Created)			
DT	01-NOV-1996	(TrEMBLrel. 01, Last sequence update)			
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)			
DE	C-Serate-1	protein (Fragment).			
OS	Gallus gallus	(Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;				
OC	Gallus.				
OX	NCBI_TaxID=9031;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE.				
RC	TISSUE=Otic explant, and Spinal cord;				

```
RX MEDLINE=96175595; PubMed=8631496; DOI=10.1006/dbio.1996.0069;
RA Myat A., Henrique D., Ish-Horowicz D., Lewis J.;
RT "A chick homologue of Serrate and its relationship with Notch and
RL Delta homologues during central neurogenesis.";
CC -I- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
DR ENBL; X95283; CAA64604.1; -; mRNA.
DR HSP; P00740; IEDM.
DR Ensemble; ENSGALG00000009020; Gallus gallus.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0007154; P:cell communication; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0007219; P:Notch signaling pathway; IEA.
DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR001774; DSL.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_YI.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR011651; MNFL.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF01414; DSL; 1.
DR Pfam; PF00008; EGF; 12.
DR Pfam; PF07645; EGF_Ca; 1.
DR Pfam; PF07657; MNFL; 1.
DR PRINTS; PR00010; EGFBL00D.
DR SMART; SM00051; DSL; 1.
DR SMART; SM00179; EGF_CA; 10.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 10.
DR PROSITE; PS01051; DSL; 1.
DR PROSITE; PS00022; EGF_1; 16.
DR PROSITE; PS01186; EGF_2; 12.
DR PROSITE; PS00026; EGF_3; 16.
DR PROSITE; PS01187; EGF_CA; 8.
KW Developmental protein; EGF-like domain; Notch signaling pathway.
FT NON_TER 1
SQ SEQUENCE 1193 AA; 131040 MW; 24E94D09B987935 CRC64;

Query Match 87.5%; Score 5469.5; DB 2; Length 1193;
Best Local Similarity 87.0%; Pred. No. 0;
Matches 907; Conservative 57; Mismatches 77; Indels 1; Gaps 1;

Qy 27 KVGASGQFELEILSMQVNGELQNGCCGARNPGDKCTRDCEDTYFKVCLKEYQSRV 86
Db 1 QVASASGQFELEILSVQVNGVLQNGCCDGRNPGDKCTRDCEDTYFKVCLKEYQSRV 60

Qy 87 TAGGPGCFGSGSTPVIQGNTEFNLKASRGNDNRIVLPFSFAWPRSYYTLLEAVDSSNDTV 146
Db 61 TAGGPGCFGSGSTPVIQGNTEFNLYSRNEKNRIVIPFSFAWPRSYYTLLEAVDNDNDST 120

Qy 147 QPDSIIKASHGMINPSRQWTLKQNTGVVAHFYQIRVTCDDYYVYFGNCKFCRPRDDF 206
Db 121 NPDRIIKASHGMINPSRQWTLKHTGAHFYQIRVTCAEHYVYFGNCKFCRPRDDF 180

Qy 207 FGHYACDQNGKTCMEGWMGPECNRAICRQCSKPKHGSCKLPDGCRCOYHQWGLYCDKCI 266
Db 181 FTHHTCDQNGKTCLEGTWGPCECNKAICRQCSKPKHGSCTVPGECCRCQYHQWGLYCDKCI 240

Qy 267 PHPCGVHGI CNEPWQCLCTETWNGQLCDKDLNYCGTHQPCCLNGCTCSNTGPDKYQCSCE 326
Db 241 PHPCGVHGTCTEPWQCLCTETWNGQLCDKDLNYCGTHPPCLNGGTCSTGPDKYQCSCE 300

Qy 327 GYSPNCEIAEHACLSDPCHNRGSKETSLEFCECSPGWTGPTCSTNIDDCSPNNCSHG 386
Db 301 GYSGNCEIAEHACLSDPCHNGSSCLEISTGFCVCAGWAGPTCTDNIDDCSPNCGHG 360

Qy 387 GTCQDLVNGFKVCPQWTKTQCLDANECEBAKPCVNAKSCNLIASYYCDCLPGWNGQN 446
Db 361 GTCQDLVDGFKCICPPQWTKTQCLDANECEBCKPKCVNANSCRNLISYCYCDCTGWSGHN 420
```

DR InterPro; IPR001774; DSL.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR011651; MNFL.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF01414; DSL; 1.
DR Pfam; PF00008; EGF; 13.
DR Pfam; PF07645; EGF_Ca; 1.
DR Pfam; PF07657; MNFL; 1.
DR PRINTS; PR00010; EGFBLD.
DR SMART; SM00051; DSL; 1.
DR SMART; SM00179; EGF_Ca; 10.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00010; ASX HYDROXYL; 10.
DR PROSITE; PS00022; EGF_1; 16.
DR PROSITE; PS01186; EGF_2; 12.
DR PROSITE; PSS0026; EGF_3; 15.
DR PROSITE; PS01187; EGF_Ca; 8.
SQ SEQUENCE 1214 AA; 133807 MW; A9C09C420CC157C2 CRC64;

Query Match 84.2%; Score 5263; DB 2; Length 1214;
Best Local Similarity 82.3%; Pred. No. 0;
Matches 879; Conservative 87; Mismatches 96; Indels 6; Gaps 4;

Qy 1 MRSPTGRSGRPSLALLCALRAKVCAGSGQFELEILSMQNVNGELQNGCCGARN 60
Db 1 MRFPRL--RALSPL-LVSALLLRLKIVSTIASGQFELEILSMQNPNGELQNGCCDQRN 56

Qy 61 PGDRKTRDRCDTYFKVCLKEYQSRVTAGPCSPGSGSTVPVIGNTNFKASRGNDNRRI 120
Db 57 PTDKCRDRCDTYFKVCLKEYQSRVSAGGACSGFTGYTPVIGNSFNKYSRNRNRRI 116

Qy 121 VLPFSFAMPASYLLVLEAWDSNDTQVDSITIEKASHGMINSPRQWOTLKONTGVVAHE 180
Db 117 VLPFSFAMPASYLLVLEAWDYNDTNDPGLIDKALHSGMINSPRQWOTLKQNGTYFE 176

Qy 181 YQIRVTCDYDYYGFCNKFCRPRDDFFGHYACDQNGKTCMEGMGPECNRAICRQCSGP 240
Db 177 YQIRVICDEHYGFCNKFCRPRDDFFGHYTCDLNGNKTCLEGMGPECSTALCRQCSS 236

Qy 241 KHGSKLPGRCRCOYGHQGLYCDKIPHPGCVHGI CNEPQCICETWGGQLCDKDLNYC 300
Db 237 KHGTCKTPGRCRCOYGHQGLYCDKIPHPGCVHGT CNEPQCICETWGGQLCDKDLNYC 296

Qy 301 GTHPQCLNGTCSNTGPDKYQCSPEGYSGPNCBIAEASHACLSDPCNHRGSKETSIGFEC 360
Db 297 GTVQPCNLNGTCSNTGPDKYQCSPEGYSGQNCBIAEASHACLSDPCNHRGSKETSIGFEC 356

Qy 361 ECSFGWTGPTCSTNIDDCSPNCSHGTCQDLVNGFKVCYCPQWTKTQCLDANECEAKP 420
Db 357 QCARGWTGPTCSINIDCSFNPCGYGTCQDLIDGFKCICPSQWTKTQCLDANECEAKP 416

Qy 421 CVNAKSKNLIASYCDLPGWQNCNDININDCLGQONDASCRDLVNGYRCICPPGYA 480
Db 417 CVNANSRNLIGSYCNCLPGWQNCNDININDCLGQONGTCKDLVNGYRCICPPGYA 476

Qy 481 GDHCERDIDRCASNPLCNGHCQNEINRFQCLPTGFSNLCOLDIDYCEPNPCONGAOC 540
Db 477 GERCEKDVNCEVSNPLCNGHCQNEINRFQCLPAGFSNLCQLDIDYCEPDPQNGAOC 536

Qy 541 YNRASDYFCPCPDYEGKNGSHLKHCRTRTPCEVIDSCITVAMASNDTPPEGVRISSNVCG 600
Db 537 FNLATDYFCNCSDEYEGKNGSHLKHCRTRTPCEVIDSCITVAMASNDTPPEGVRISSNVCG 596

Qy 601 PHGKCKSQSGKFTCDCKNGFTGTGYCHENINDCESNPRNGTICDGVNSYKICISDQWE 660
Db 597 PHGKCRSQSGKFTCECKGFTGTGYCHENINDCESNPNNGGTCIDGVNSYKICISNGWE 656

Qy 661 GAYCETNINDCONPCNNGTCDRLVNDYFCDCKNGWKGTCHSRSDCEATCNNGGTC 720
Db 657 GIYCETNINDCSNPNNGTCDRLVNDYFCECKNGWKGTCHSRSDCEATCNNGGTC 716

Qy 721 YDEGDAFKMCPGWEGETTCNIARNSSCLNPNCHNGTGVVNGBSFTCVCKEGWEGPICA 780
Db 717 YDEGDTFKCICSPWEGATCNIAARNSSCLNPNCFNGGTCTVVRGDSFTCVCKEGWEGPTCS 776

Qy 781 QNTNDCSPHPCYNSGTCVGDGNMYRCACAGFAGPDCRININECQSSPCAFGATCVDBIN 840
Db 777 QNTNDCSPHPCYNSGTCVGDGNMYRCACAGFAGPDCRININECQSSPCAFGATCIDBIN 836

Qy 841 GYRCVCPPHGSGAKCQEVSGRPCITMGSVIPDGAKNDDDCNTCQCLNGRIACSKVWCGPR 900
Db 837 GYRCVCPPHGSGRPCQEVTCGRPCITMGHMPDGEKWNDCNQCLNGKVKTCVKWCGPQ 896

Qy 901 PCLLKHGSECPGQSCIPILDDQCFVHPCTGVGECRSSLSLOPVTIKTSDSYQD-NCA 959
Db 897 RCDIN-GDSECPAGQTCVPRDNCFCVPPCTGLGECWFTNPQPPVTKTCNANASYQDASCA 955

Qy 960 NITFTFNKEMWSPGLTTEHICSELNINILKUNSAEYSIYIACEPSPSANNEIHVAISAE 1019
Db 956 NITFTFNKEMWSPGLTTESICNELNINILKUNSAEYSIYTCLSPSASNEIHVAISVE 1015

Qy 1020 DIRDDGNPIKEITDKIIDLVSKRDGNSSLIAAFAEVRVQRPLKNRTD 1067
Db 1016 ENGGRNPVKDITERIIDLVSKRDANNTFTIASVAEVRVQTRHTYKNGTD 1063

RESULT 7
Q5U4U1 XENLA
ID Q5U4U1 XENLA PRELIMINARY; PRT; 1051 AA.
AC OSU4U1
DT 01-FEB-2005 (TEMBLrel. 29, Created)
DT 01-FEB-2005 (TEMBLrel. 29, Last sequence update)
DE 01-FEB-2005 (TEMBLrel. 29, Last annotation update)
DE X-Serrate-1 protein.
GN Name=X-Serrate-1;
OS Xenopus laevis (African clawed frog).
OC Sauravata; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Oocytes;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative";
RL Dev. Dyn. 225:384-391 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Oocytes;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udutin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

FT	DISULFID	342	357	By similarity.
FT	DISULFID	359	368	By similarity.
FT	DISULFID	375	386	By similarity.
FT	DISULFID	380	395	By similarity.
FT	DISULFID	397	406	By similarity.
FT	DISULFID	413	424	By similarity.
FT	DISULFID	418	433	By similarity.
FT	DISULFID	435	444	By similarity.
FT	DISULFID	451	461	By similarity.
FT	DISULFID	455	470	By similarity.
FT	DISULFID	472	481	By similarity.
FT	DISULFID	488	499	By similarity.
FT	DISULFID	493	508	By similarity.
FT	DISULFID	510	519	By similarity.
FT	DISULFID	526	537	By similarity.
FT	DISULFID	531	546	By similarity.
FT	DISULFID	548	557	By similarity.
FT	DISULFID	630	641	By similarity.
FT	DISULFID	635	650	By similarity.
FT	DISULFID	652	661	By similarity.
FT	DISULFID	668	679	By similarity.
FT	DISULFID	673	688	By similarity.
FT	DISULFID	690	699	By similarity.
FT	DISULFID	706	717	By similarity.
FT	DISULFID	711	726	By similarity.
FT	DISULFID	728	737	By similarity.
FT	DISULFID	783	794	By similarity.
FT	DISULFID	788	803	By similarity.
FT	DISULFID	805	814	By similarity.
FT	DISULFID	821	832	By similarity.
FT	DISULFID	826	841	By similarity.
FT	DISULFID	843	852	By similarity.
SQ	SEQUENCE	1213 AA; 133366 MW; 5CSF16A7E20D9534	CRC64;	
Query Match				
Best Local Similarity 76.3%; Score 4768.5; DB 1; Length 1213;				
Matches 789; Conservative 108; Mismatches 151; Indels 5; Gaps 5;				
Qy	17	LLALLCALRAKVCASGQFELEILLSMQNVANGELQNGNCCGGARNPGDKRKTRECDTYFK	76	
Db	14	LHAFLLCLRTVSDASGHFELEILLSMQNANGELQNGACCDGARNPADRKCTRECDTYFK	73	
Qy	77	VCLKEYQSRVTAGPCFSFGSGSTPVI GGNTFNLKASGRNDRNRIVLPPSFAPWPSYTLIV	136	
Db	74	VCLKEYQSRVSSAGACSFSGTSTPVLGNKPFSTKGTR -SEKSRIVLPPSFAPWPSYTLIV	132	
Qy	137	BANDSSNDTV -QPDSSIIEKASHSGWNPISROWQTLKONTVAHPEYQIRVTCDDYVYGF	195	
Db	133	EALDFNNETASBEGKLEKAYHSGMINPNRQWRLTHNGPVAQFEYQIRVTCLEHYVGF	192	
Qy	196	CNKFCDPRDDFFGHYACDQNGKTCMEGWMGPECNRAICROCCSPKHGSKLPGDCRCQY	255	
Db	193	CNKFCDPRDFFGHYTCDDQNGKTCLESGWTGPDNTAI CROCCSTEAGSKQPGCKCLY	252	
Qy	256	GWGLYCDKCI PHPGCVHGI CNBPWCCLCETNWWGQLCDKDLNYCGTHQPCLINGTCSNT	315	
Db	253	GWGPGYCDKCI PHPGCVHGT CVBPWPQCLCDTNWWGQLCDKDLNYCGTHQPCLINGTCSNT	312	
Qy	316	GPDKYQCSCEPGYSGNPCEIAEHAACLSDPCHNRGSKETS LGFPCECSPGHTGTPCSTNI	375	
Db	313	GPDKYQCSCEBDYSGVNCBBAEHAACLSPNCANGGTCKETS SGYEGCHCAIGWSGTSCEINV	372	
Qy	376	DDCSPNNSHGCTCDLVNGFKVCVCPQWTGKTQLDANECBAKPCVNAKSKCNLIASY	435	
Db	373	DDCTPNQCKIGTCDLVNGFKCACP PHTWTGKTQIDANECEDKPCVNAKSKCHNLIGAYF	432	
Qy	436	CDCLPGWMGQNC DININDCLGQCNQDASCRDLVNGYRCI PPGYAGDHCERIDECASNP	495	
Db	433	CECLPGWSGQNC DININDCKGQCLNGGTCCKDLVNGYRCLCPPGYTGBQCEKDVDECASSP	492	
Qy	496	CLNGHGHCONENRFOCLCPTGFGSGLVQLDLDIYCEPMPQNGAOCYNPASDYFKCKPDDY	555	
Db	493	CLNGGRQDSEVNGPQCLCPAGFSQCLQCLTDIDYCKPNPCQNGAQCFNLASDYFKCKPDDY	552	

Qy	556	EGKNC	SHLKHCHRTTPCEVIDDSCTVAMASNDTPEGVYIISNVCGPHGCKCSOGGKFTC	615
Db	553	EGKNC	SHLKHCHRTTSCQVIDDSCTVAVASNTPEGVYIISNVCGPHGRCSRAGGQFTC	612
Qy	616	DCNKG	FGTGYCHENINDCESNPRNGGTCIDGVNSYKICISDGEWEGAYCETNINDCSQNP	675
Db	613	ECQEG	FRGTGYCHENINDCESNPRNGGTCIDKVNYYQICADGWEGVHCEINIDDCSLNP	672
Qy	676	CHNGGT	CRDLVNDPYCDCKNGWKTKCHSRDSQCDDEATCNNGGTCYVDGDAFKCMCPGGW	735
Db	673	CLNKG	ACODLVNDPYCBRNNGWKTKCHSRDSQCDDEATCNNGGTCYVDGDTFKCRSPGW	732
Qy	736	EGTTC	NIAKSSCLPNCHNGGTCYVNGESFTCVCKEGWEGPICAQNTWDCSHPCCVNSG	795
Db	733	EGATC	NIAKSSCLPNCHNGGTCYVNGDSFNCVCKEGWEGSTENTNDCNPHPCVNSG	792
Qy	796	TCVDG	DNWYCECAPGAGPDCRININECOSSPCAFATCVDINGYRCVCPGPHSGAKC	855
Db	793	TCVDG	DNWYCECAPGAGPDCRININECOSSPCAFGSTCVDINGYRCVCPGRIGPDC	852
Qy	856	QEVSG	RCITWGSVIPDGAKNWDDCNTCQCLNGRIACSWKGPRLCLLKHGSHSECPGQ	915
Db	853	QEVVG	RCIANGQVTADGAKWEEDCNTCQONGRIHCTMWCCKPSCRIKARGGCPASQ	912
Qy	916	SCPI	LDQCFVHCTGVGECRSSLOPVKTC-TSDSYQDNCANITFTFNKEMKSPGL	974
Db	913	SCVPI	KEQCFVHCTGVGECRSSLOPVKTC-TSDSYQDNCANITFTFNKEMKSPGL	971
Qy	975	TTEHC	SELNLTNLTKNVSAEYSIYIACEPSPSANNHVAISAEIRDGNPKETDK	1034
Db	972	SVEHC	NELHWHYLLKNLTSTAYVSIACEPSSASNIHISISTEETDRSPKIDITVQ	1031
Qy	1035	IIDL	VSKRDNSSIIAAVAEVRORRPLKNRTD	1067
Db	1032	IIDL	VSKHNGNSTIIKAITGVRVHOIP-SPKTD	1063
RESULT 9				
JAGIA	BRARE	STANDARD;	PRT;	1242 AA.
ID	JAGIA_BRARE	Q90Y57;		
AC	Q90Y57;			
DT	28-FEB-2003	(Rel. 41, Created)		
DT	28-FEB-2003	(Rel. 41, Last sequence update)		
DT	13-SEP-2005	(Rel. 48, Last annotation update)		
DE	Jagged-1a precursor (Jagged1a) (Jagged1).			
GN	Name=Jag1a; Synonyms=jag1;			
OS	Brachydanio rerio (Zebrafish) (Danio rerio).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;			
OC	Cyprinidae; Danio.			
OX	NCBI_TaxID=7955;			
RN	(1)			
RP	NUCLEOTIDE SEQUENCE.			
PA	Oda T., Chandrasekharappa S.C.;			
RT	"Isolation, characterization and expression analysis of zebrafish			
RT	Jagged genes."			
RL	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.			
CC	-1- FUNCTION: Ligand for multiple Notch receptors and involved in the			
CC	mediation of Notch signaling (By similarity). Seems to be involved			
CC	in cell-fate decisions.			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).			
CC	-1- SIMILARITY: Contains 1 DSL domain.			
CC	-1- SIMILARITY: Contains 17 EGF-like domains.			
CC	-----			
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use as long as its content is in no way modified and this statement is not			
CC	removed.			
CC	-----			
DR	EMBL; AF229448; AAL08213.1; -; mRNA.			
DR	HSP; P01132; 1GK5.			

DR	ZFIN; ZDB-GENE-011128-2; jag1a.			
DR	InterPro; IPR000152; Asx_hydroxyl_s.			
DR	InterPro; IPR001774; DSL.			
DR	InterPro; IPR000742; EGF_2.			
DR	InterPro; IPR001881; EGF_Ca_bd.			
DR	InterPro; IPR001438; EGF_II.			
DR	InterPro; IPR006209; EGF_like.			
DR	InterPro; IPR011651; MNFL.			
DR	InterPro; IPR001007; VWF_C.			
DR	Pfam; PF01414; DSL; 1.			
DR	Pfam; PF00008; EGF; 13.			
DR	Pfam; PF07657; MNFL; 1.			
DR	PRINTS; PR00010; EGFBL00D.			
DR	SMART; SM00051; DSL; 1.			
DR	SMART; SM00179; EGF_CA; 9.			
DR	SMART; SM00214; VMC; 1.			
DR	PROSITE; PS00010; ASX_HYDROXYL; 9.			
DR	PROSITE; PS1051; DSL; 1.			
DR	PROSITE; PS00022; EGF_1; 16.			
DR	PROSITE; PS01186; EGF_2; 12.			
DR	PROSITE; PS00026; EGF_3; 15.			
DR	PROSITE; PS01187; EGF_CA; 8.			
KW	Calcium; Developmental protein; EGF-like domain; Glycoprotein;			
KW	Notch signaling pathway; Repeat; Signal; Transmembrane.			
FT	SIGNAL	1	28	Potential.
FT	CHAIN	29	1242	Jagged-1a.
FT	TOPO_DOM	29	1070	Extracellular (Potential).
FT	TRANSMEM	1071	1095	Potential.
FT	TOPO_DOM	1096	1242	Cytoplasmic (Potential).
FT	DOMAIN	186	230	DSL.
FT	DOMAIN	231	264	EGF-like 1.
FT	DOMAIN	262	295	EGF-like 2.
FT	DOMAIN	297	335	EGF-like 3.
FT	DOMAIN	337	373	EGF-like 4.
FT	DOMAIN	375	411	EGF-like 5;
FT	DOMAIN	413	449	calcium-binding (Potential).
FT	DOMAIN	451	486	EGF-like 6;
FT	DOMAIN	488	524	calcium-binding (Potential).
FT	DOMAIN	526	562	EGF-like 7;
FT	DOMAIN	575	630	calcium-binding (Potential).
FT	DOMAIN	632	668	EGF-like 8;
FT	DOMAIN	670	706	EGF-like 9.
FT	DOMAIN	708	744	EGF-like 10.
FT	DOMAIN	747	783	EGF-like 11; calcium-binding (Potential).
FT	DOMAIN	785	821	EGF-like 12; calcium-binding (Potential).
FT	DOMAIN	823	859	EGF-like 13.
FT	DOMAIN	917	959	EGF-like 14.
FT	CARBOHYD	141	141	EGF-like 15; calcium-binding (Potential).
FT	CARBOHYD	218	218	EGF-like 16; calcium-binding (Potential).
FT	CARBOHYD	385	385	EGF-like 17.
FT	CARBOHYD	560	560	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	748	748	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	960	960	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	991	991	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1046	1046	N-linked (GlcNAc. . .) (Potential).
FT	DISULFID	235	246	By similarity.
FT	DISULFID	239	252	By similarity.
FT	DISULFID	254	263	By similarity.
FT	DISULFID	266	277	By similarity.
FT	DISULFID	272	283	By similarity.
FT	DISULFID	285	294	By similarity.
FT	DISULFID	301	313	By similarity.
FT	DISULFID	307	323	By similarity.
FT	DISULFID	325	334	By similarity.
FT	DISULFID	341	352	By similarity.
FT	DISULFID	346	361	By similarity.
FT	DISULFID	363	372	By similarity.
FT	DISULFID	379	390	By similarity.
FT	DISULFID	384	399	By similarity.
FT	DISULFID	401	410	By similarity.
FT	DISULFID	417	428	By similarity.
FT	DISULFID	422	437	By similarity.
FT	DISULFID	439	448	By similarity.

FT	DISULFID	455	465	By similarity.	603	GKCKSOS--GGKFTCDNKGFTGTYCHENINDCESNPRNGGTCIDGVNSYKICISDGWE	660
FT	DISULFID	459	474	By similarity.	604	GRCRSHAGGHFSCDQDGFYTYCHENINDCESSCLSGGTCDKINAYQCICADGWE	663
FT	DISULFID	476	485	By similarity.	661	GAYCETNINDCSNPCHNGGTCRDLYNDVFCDCNKGKGTCHSRDSQCDEATCNGGTC	720
FT	DISULFID	492	503	By similarity.	664	GPNCETNIDCRTNPCRDRGVCRDLYNDVFCENCGWKGTCHSRSCQDEDFCNGGTC	723
FT	DISULFID	497	512	By similarity.	721	YDEGDAFKCPCPGWEGTTCNTIARNSSCLPNPCHNGGTCVYNGESFTVCCKEGWEGPICA	780
FT	DISULFID	514	523	By similarity.	724	SDEGDSFKCLCSPGWEGATCNIAKNSCLPNPCENGATCVVTDGFTVCCKEGWEGPICS	783
FT	DISULFID	523	541	By similarity.	781	QNTNDCSPHCYNSTGTCVGDNDWYRCCECAPGAGPAGPCRNINECQSSPCAFGATCVDDEIN	840
FT	DISULFID	530	541	By similarity.	784	QNSNDCNPQPCYNSTGTCVGDNDWYRCCECAPGAGPAGPCRNINECQSSPCAFGATCVDDEIN	843
FT	DISULFID	535	550	By similarity.	841	GYRCVCPGHSGAKCOEVSRCPTMGSVIPDCAKWDCCNTCCQCLNGRTACSKVWCGPR	900
FT	DISULFID	552	561	By similarity.	844	GYRCLCPPERTGPRCOEVTGRCVIGRTAVDGAKEADCNCTCYCHKGIYVCTKLCFGPK	903
FT	DISULFID	561	572	By similarity.	901	PC-LLHKGHSECPSSGQSCIPILDQCFVHPCTGVGECRSSSLQPVTKTCTSDSYQDNCA	959
FT	DISULFID	566	572	By similarity.	904	ACRMLGSGRGDCPTGQLCVPRDEQCFVKPCSSQSGECWSAHRPAVRTHCQPDHSH---CA	959
FT	DISULFID	567	585	By similarity.	960	NITFTFNKEMSPGLTTEHICSELNMLNKLKNSVSIYIACEPSPSANNEIHVAISAE	1019
FT	DISULFID	571	582	By similarity.	960	NVTFTFNKDTMPQGVTVQVCRHLHYTKNTVSFISVSCSELSSAASNEIHVAHVT	1019
FT	DISULFID	573	587	By similarity.	1020	DIRDDGN-PIKETDKIIDLKSKRDNSSILIAVAEVRVORRPLKN	1064
FT	DISULFID	574	588	By similarity.	1020	ENGIHRVFPVKETDNIIDLVSKHSANSSVIGSIAEVRVORRQKQPN	1065
FT	DISULFID	575	588	By similarity.			
FT	DISULFID	576	588	By similarity.			
FT	DISULFID	577	588	By similarity.			
FT	DISULFID	578	588	By similarity.			
FT	DISULFID	579	588	By similarity.			
FT	DISULFID	580	588	By similarity.			
FT	DISULFID	581	588	By similarity.			
FT	DISULFID	582	588	By similarity.			
FT	DISULFID	583	588	By similarity.			
FT	DISULFID	584	588	By similarity.			
FT	DISULFID	585	588	By similarity.			
FT	DISULFID	586	588	By similarity.			
FT	DISULFID	587	588	By similarity.			
FT	DISULFID	588	588	By similarity.			
FT	DISULFID	589	588	By similarity.			
FT	DISULFID	590	588	By similarity.			
FT	DISULFID	591	588	By similarity.			
FT	DISULFID	592	588	By similarity.			
FT	DISULFID	593	588	By similarity.			
FT	DISULFID	594	588	By similarity.			
FT	DISULFID	595	588	By similarity.			
FT	DISULFID	596	588	By similarity.			
FT	DISULFID	597	588	By similarity.			
FT	DISULFID	598	588	By similarity.			
FT	DISULFID	599	588	By similarity.			
FT	DISULFID	600	588	By similarity.			
FT	DISULFID	601	588	By similarity.			
FT	DISULFID	602	588	By similarity.			
FT	DISULFID	603	588	By similarity.			
FT	DISULFID	604	588	By similarity.			
FT	DISULFID	605	588	By similarity.			
FT	DISULFID	606	588	By similarity.			
FT	DISULFID	607	588	By similarity.			
FT	DISULFID	608	588	By similarity.			
FT	DISULFID	609	588	By similarity.			
FT	DISULFID	610	588	By similarity.			
FT	DISULFID	611	588	By similarity.			
FT	DISULFID	612	588	By similarity.			
FT	DISULFID	613	588	By similarity.			
FT	DISULFID	614	588	By similarity.			
FT	DISULFID	615	588	By similarity.			
FT	DISULFID	616	588	By similarity.			
FT	DISULFID	617	588	By similarity.			
FT	DISULFID	618	588	By similarity.			
FT	DISULFID	619	588	By similarity.			
FT	DISULFID	620	588	By similarity.			
FT	DISULFID	621	588	By similarity.			
FT	DISULFID	622	588	By similarity.			
FT	DISULFID	623	588	By similarity.			
FT	DISULFID	624	588	By similarity.			
FT	DISULFID	625	588	By similarity.			
FT	DISULFID	626	588	By similarity.			
FT	DISULFID	627	588	By similarity.			
FT	DISULFID	628	588	By similarity.			
FT	DISULFID	629	588	By similarity.			
FT	DISULFID	630	588	By similarity.			
FT	DISULFID	631	588	By similarity.			
FT	DISULFID	632	588	By similarity.			
FT	DISULFID	633	588	By similarity.			
FT	DISULFID	634	588	By similarity.			
FT	DISULFID	635	588	By similarity.			
FT	DISULFID	636	588	By similarity.			
FT	DISULFID	637	588	By similarity.			
FT	DISULFID	638	588	By similarity.			
FT	DISULFID	639	588	By similarity.			
FT	DISULFID	640	588	By similarity.			
FT	DISULFID	641	588	By similarity.			
FT	DISULFID	642	588	By similarity.			
FT	DISULFID	643	588	By similarity.			
FT	DISULFID	644	588	By similarity.			
FT	DISULFID	645	588	By similarity.			
FT	DISULFID	646	588	By similarity.			
FT	DISULFID	647	588	By similarity.			
FT	DISULFID	648	588	By similarity.			
FT	DISULFID	649	588	By similarity.			
FT	DISULFID	650	588	By similarity.			
FT	DISULFID	651	588	By similarity.			
FT	DISULFID	652	588	By similarity.			
FT	DISULFID	653	588	By similarity.			
FT	DISULFID	654	588	By similarity.			
FT	DISULFID	655	588	By similarity.			
FT	DISULFID	656	588	By similarity.			
FT	DISULFID	657	588	By similarity.			
FT	DISULFID	658	588	By similarity.			
FT	DISULFID	659	588	By similarity.			
FT	DISULFID	660	588	By similarity.			
FT	DISULFID	661	588	By similarity.			
FT	DISULFID	662	588	By similarity.			
FT	DISULFID	663	588	By similarity.			
FT	DISULFID	664	588	By similarity.			
FT	DISULFID	665	588	By similarity.			
FT	DISULFID	666	588	By similarity.			
FT	DISULFID	667	588	By similarity.			
FT	DISULFID	668	588	By similarity.			
FT	DISULFID	669	588	By similarity.			
FT	DISULFID	670	588	By similarity.			
FT	DISULFID	671	588	By similarity.			
FT	DISULFID	672	588	By similarity.			
FT	DISULFID	673	588	By similarity.			
FT	DISULFID	674	588	By similarity.			
FT	DISULFID	675	588	By similarity.			
FT	DISULFID	676	588	By similarity.			
FT	DISULFID	677	588	By similarity.			
FT	DISULFID	678	588	By similarity.			
FT	DISULFID	679	588	By similarity.			
FT	DISULFID	680	588	By similarity.			
FT	DISULFID	681	588	By similarity.			
FT	DISULFID	682	588	By similarity.			
FT	DISULFID	683	588	By similarity.			
FT	DISULFID	684	588	By similarity.			
FT	DISULFID	685	588	By similarity.			
FT	DISULFID	686	588	By similarity.			
FT	DISULFID	687	588	By similarity.			
FT	DISULFID	688	588	By similarity.			
FT	DISULFID	689	588	By similarity.			
FT	DISULFID	690	588	By similarity.			
FT	DISULFID	691	588	By similarity.			
FT	DISULFID	692	588	By similarity.			
FT	DISULFID	693	588	By similarity.			
FT	DISULFID	694	588	By similarity.			
FT	DISULFID	695	588	By similarity.			
FT	DISULFID	696	588	By similarity.			
FT	DISULFID	697	588	By similarity.			
FT	DISULFID	698	588	By similarity.			
FT	DISULFID	699	588	By similarity.			
FT	DISULFID	700	588	By similarity.			
FT	DISULFID	701	588	By similarity.			
FT	DISULFID	702	588	By similarity.			
FT	DISULFID	703	588	By similarity.			
FT	DISULFID	704	588	By similarity.			
FT	DISULFID	705	588	By similarity.			
FT	DISULFID	706	588	By similarity.			
FT	DISULFID	707	588	By similarity.			
FT	DISULFID	708	588	By similarity.			
FT	DISULFID	709	588	By similarity.			
FT	DISULFID	710	588	By similarity.			
FT	DISULFID	711	588	By similarity.			
FT	DISULFID	712	588	By similarity.			
FT	DISULFID	713	588	By similarity.			
FT	DISULFID	714	588	By similarity.			
FT	DISULFID	715	588	By similarity.			
FT	DISULFID	716	588	By similarity.			
FT	DISULFID	717	588	By similarity.			
FT	DISULFID	718	588	By similarity.			
FT	DISULFID	719	588	By similarity.			
FT	DISULFID	720	588	By similarity.			
FT	DISULFID	721	588	By similarity.			
FT	DISULFID	722	588	By similarity.			
FT	DISULFID	723	588	By similarity.			
FT	DISULFID	724	588	By similarity.			
FT	DISULFID	725	588	By similarity.			
FT	DISULFID	726	588	By similarity.			
FT	DISULFID	727	588	By similarity.			
FT	DISULFID	728	588	By similarity.			
FT	DISULFID	729	588	By similarity.			
FT	DISULFID	730	588	By similarity.			
FT	DISULFID	731	588	By similarity.			
FT	DISULFID	732	588	By similarity.			
FT	DISULFID	733	588	By similarity.			
FT	DISULFID	734	588	By similarity.			
FT	DISULFID	735	588	By similarity.			
FT	DISULFID	736	588	By similarity.			
FT	DISULFID	737	588	By similarity.			
FT	DISULFID	738	588	By similarity.			
FT	DISULFID	739	588	By similarity.			
FT	DISULFID	740	588	By similarity.			
FT	DISULFID	741	588	By similarity.			
FT	DISULFID	742	588	By similarity.			
FT	DISULFID	743	588	By similarity.			
FT	DISULFID	744	588	By similarity.			
FT	DISULFID	745	588	By similarity.			
FT	DISULFID	746	588	By similarity.			
FT	DISULFID	747	588	By similarity.			
FT	DISULFID	748	588	By similarity.			
FT	DISULFID	749	588	By similarity.			
FT	DISULFID	750	588	By similarity.			
FT	DISULFID	751	588	By similarity.			
FT	DISULFID	752	588	By similarity.			
FT	DISULFID	753	588	By similarity.			
FT	DISULFID	754	588	By similarity.			
FT	DISULFID	755	588	By similarity.			
FT	DISULFID	756	588	By similarity.			
FT	DISULFID	757	588	By similarity.			
FT	DISULFID	758	588	By similarity.			
FT	DISULFID	759	588	By similarity.			
FT	DISULFID	760	588	By similarity.			
FT	DISULFID	761	588	By similarity.			
FT	DISULFID	762	588	By similarity.			
FT	DISULFID	763	588	By similarity.			
FT	DISULFID	764	588	By similarity.			
FT	DISULFID	765	588	By similarity.			
FT	DISULFID	766	588	By similarity.			
FT	DISULFID	767	588	By similarity.			
FT	DISULFID	768	588	By similarity.			
FT	DISULFID	769	588	By similarity.			
FT	DISULFID	770	588	By similarity.			
FT	DISULFID	771	588	By similarity.			
FT	DISULFID	772	588	By similarity.			
FT	DISULFID	77					

SQL	SEQUENCE	1364 AA; 148433 MW; 8E0ACF1A3BF1BC7BE CRC64;
	Query Match	68.7%; Score 4293.5; DB 2; Length 1364;
	Best Local Similarity	62.3%; Pred. No. 6.9e-283;
	Matches	745; Conservative 120; Mismatches 177; Indels 155; Gaps 14;
QY	15	SLLLALLCALRAKVCGASGOFELIILSMQNVGELQNGCCGARNPDGRKCTRDECDTV 74
DB	7	SLVAALLLLCQAQVSEGSQFELQILSMHNVGELLSSVCCDGAARSAGDKRCARDECATP 66
QY	75	FKVCLKEYQSRVTAGGSCSPGSGSTPVI GGNFTNLKASRGNDNRIVLPSPFAWPSRYTL 134
DB	67	FKVCLKEYQSRVTHAAGPCSPGMGSTPVLGGNTFSLKSVRIDKSRITLTPSPFAWPSRYTL 126
QY	135	LVEAMDSNDTVQDPS-IIEKASHSGMINPSROWQTLKONTGVAHFQYQIRVTCDDYYG 193
DB	127	IVEAFDSNETSGADSRLLIEKAPHSGMINPSPOWKLTHTNGPVAQFYQIRVTCDEHYG 186
QY	194	FGCNKFCRPDRDDFFGHYACDQNGKTCMEGWMGPECNRAICROGCSPKHSGCKLPGDCR- 252
DB	187	FGCNKFCRPDRDEPFGHYTCDYGNGKTCLEGSGPDCNTAICROGCSIEHSGCKEPGBCNV 246
QY	253	CQYGWQ-----GLYCDKC-----IHPGCVHGIC-----NEPQCL----- 283
DB	247	CAMIRQIMSRRAARSCLGRQDAANPPASPSPRWIWNRCVSPQFQSLGAQAWLLTHTGN 306
QY	284	-----CBTNWGG-----QLCD- 294
DB	307	QVVCMAGRASTVTSASPTPAAYMAPVWSPGSA CATPTGAATFAKEPPVLCANRLKLCGT 366
QY	295	-KOLNYGCTHQPLNGGTCSTNGTDPKYQCSPEYSGPNCIEIAHACLSDPCHNRGSCKE 353
DB	367	LPDLNYGCTHQPLNGGTCINTGDPKYQCTCAEGYSGANCERAEHACLSPGCSNGGSCSE 426
QY	354	TSLGFECECSPGWTPGTCSTNIDDCSPNNCSHGCTCODLVNGFKVCPPQWTKTC----- 409
DB	427	TSQGYEQCAAGWSGSPCTIIDDCAPNPNCHGGTCQDLVNGYKCHCPPQWMTKTLIGE 486
QY	410	-----QLDANECEAKPCVNKSCNLTASYCDCLPG 441
DB	487	AASAAPHNRQLHSQMPFSPPPSPPLLDANECDSKPCVNANSCNLTGGYPCPCVP 546
QY	442	WMQNCNDINTDCLQGCQNDA SCRDLVNGYRCICPPGYAGDHCBERRIDECASNPLCNGGH 501
DB	547	WTGQNCDDINDCRDQCGGTCKDLVNGYRCMCPAGFSGEHCEKDVDECLSSPCLNGGR 606
QY	502	CQNEINPQCLPTGFGSNLCQLDIDYCEPNPQNGAQCVNRA SDYFCCKPEDYEGKNCS 561
DB	607	CQDEVNGFQCLCLAGFSGNLCQLDIDYCSNPCLNGASCNFNLATDYTCACPEDYEGKNCS 666
QY	562	HLKDHCRTTPEVITDSCTVAMASNDTPGVRYIISNVCGPHGKCKOSGCKFTCDCKNGF 621
DB	667	HLKDHCRTTTCVKVLDSCTVAVASNTPGGERYIISNVCGPHGRCKRQAGGQFTCEQEGF 726
QY	622	TGTYCHENINDCESNCRNGGTCIDGVNSYKICISDGMEGAYCEBTNINDCSQNCHNGGT 681
DB	727	RGTYCHENINDCESNPNCHNGGTCIDKVSVYQCIADGWEGDHQCLNIDDCSTSPCYNGGT 786
QY	682	CRDLVNDPFYCDCKXGWKGTCHSRDSCDEATCNNGGTCYDEGDAFKMCPGGWEGTTCN 741
DB	787	CRDLVTDFFCFCKXGWKGTCHSRRESQDEATCNNGGTCYDEGDAFQCKSPGWEGTTCN 846
QY	742	IARNSSCLPNPCHNGGTCVNNGESFTCVCKEGWEGPTCAONTNDCSHPH----- 790
DB	847	IARNSSCLPNPCENGGTCWAGESFTCVCKEGWEGPTCTQNTNDCSHPHWSALTVSAGK 906
QY	791	-----CYNSTGTCVDGDNWNRCECAPGAGPDCRININEC 824
DB	907	ACARPPCHSRDSVKLTSELAGLRWAS IYNSGTCTVDGDNWYRCECAPGAGPDCRININEC 966
QY	825	QSSPCAFGATCVDENGYRCVCPFGHSGAKCOEYVSGRPCTITMGSVIPDGAKKWDDCMTCQ 884
DB	967	QSSPCALGSTCVDENGYRCLCPDRTGSHCHEYWRKPCFVNGHITPDGVKWDSDCMTCQ 1026

Matches	655; Conservative	137; Mismatches	238; Indels	18; Gaps	8
Qy	27	KVCASGGQFELLEILSMQNVGELQNGNCCGGARNPGDKTRDECDTYFKVCKLEYQSRV	86		
Db	1	QVSRGTGVFELQLNSVRNVNGELLNGECCDGERNPPDRDCGRDECDTYVKVCKLEYQAKI	60		
Qy	87	TAGSPCSFGSGSTPVI CGNTFNILKA-----SRGNDRNRIVLPSPFAPWPRSYTYLLVRAWD	140		
Db	61	SPGPGCSYSGSSTVFLGNLILYLANGDKYHPRGSRPETGRIVIPFYQAMPRSFTLILEAWD	120		
Qy	141	SSNDTVQ-PSDII EKASHSGMINPSROWOTLKQNTGVAHPEVQIRVTCDDYVYGCGNKF	199		
Db	121	WDNDTKSGEDLLIERVAHAGMINPEDRWTKLQFNGPVANFEVQIRVKCDENIYALCNKF	180		
Qy	200	CRPRDDFFGHYACDQNGNKTCMEGWMPGECNRAICROGCSPKHGSCKLPGDCRCQYGNQG	259		
Db	181	CGPRDDFVGHYTCDQNGKAKMEGWMBGECKQAQVCKQGCNLLHGGCSVPGCEKCHYGQG	240		
Qy	260	LYCDKCIPIHPCGVHGINEPHOCLCETNWGGQLCDKOLNYCGTHQPCPLNGGTCNTGPDK	319		
Db	241	QYCDCEVRYPGCAHSGCNEPQCNETNWGLLCNKOLNYCGNHHPCLNGGTCMNTPEDE	300		
Qy	320	YQSCSPRGYSNPCEIAEHAACLSPCHNRGSKETSLSGFCECSPGWTGPTCSNTIDDCS	379		
Db	301	YRCACPDGYSKNCIEIAEHACVSNPCANGTTCHEISSFKCHCPSGSGPTCAIDIDECA	360		
Qy	380	PNNCSHGTCQDLVNGFKCVCPPOWTGKTCOLDANECEAKPCVNAKSCKNLIASYYDCCL	439		
Db	361	SNPCAOGGTCIDHINSFECIPCQWIGATCOLDANECEGKPCVNAYSCKNLIGGYCDCI	420		
Qy	440	PGMWGQNCININDCLGQCNDAACRDLVNGYRCICPPGYAGDHCERIDIDECASNPCPLNG	499		
Db	421	PGWKGVCHININDCHGQCHGGTCKDEVNDYHCICPRGFTGKNCEIETNECESNPQNG	480		
Qy	500	GHCQNEINRFQCLPTGFSGNLCOLDIDYCEPNPCQNGAOCYNRASDYFCCKCPEDYEGKN	559		
Db	481	GRCKDLVNGFTCLCAQGGFSGVFCBMDIDFCBNPCQNGAKCYDLGGDYCYACPDYDGN	540		
Qy	560	CSHLKDRCTTPCEVIDSCTVMAASNDTPEGVRVYISSNVCPGHCKSKSQSGKFTCDCKN	619		
Db	541	CSHLKDHCKKNSCKVIDSCTIEVSNYTTQEGIRFISNVCPGHRCLISQPGGNFTCADR	600		
Qy	620	GFTGTYCHENINDCESPNRNGGTCIDGVNSYKICICSDGWEGAYCETNINDCSQNPCHNG	679		
Db	601	GFTGIYCHENINDCLGKPCKNGGTCIDEDVSFKFCSSGMBEGELCDTNFNDCSPNPCHNG	660		
Qy	680	GTCEDLVNDYFCDCKNGWKGTCHSRDUSQDDEATCNNGGTCYDGDGAFKCMCPGGWGGT	739		
Db	661	GRCIDLVDYFCECKNDWKGTCHSRREYQCDANTCSNGGTCYDDGDTFHCSCPPFWTGST	720		
Qy	740	CNTARNSSCLPNPCHNGTCVVGNGESFTVCCKEGWGPICAOINTNDCSPHPCYNSGTCVD	799		
Db	721	CNTAKNSCTIPNCPNGGTCVSGSDSPFCICEGWEGRTCTQNTINDCNPHPCYNGGICVD	780		
Qy	800	GDWNYRCECAPGFPAGPCRININECQSSPCAFGATCTVDEINGYRCVCPPGHSGAKCOEV-	858		
Db	781	GVNWFRCECAPGFPAGPCRINIDECQSSPCGYGATCIDEINGYRCTCPPRGVRGPRQCEVI	840		
Qy	859	-SRPCITMGSVIPDGAKWDDDCNTCOCLNGRIACSKVMCGPRCLLHK-----GHSBCPS	913		
Db	841	GIGKPCWLKGMTFPHGSRWDQECNSCHLDRIDCTKVMCGKPKCLLHKHWDNSNYQCPM	900		
Qy	914	GQSCIPILDQCFVHPCTGVGCECHSSSIQPVKTKTSDSYQQDN-CANITFTFNKEMSP	972		
Db	901	GQECQEKY-MKCFHPPCTEWGECASBP LANIIRLPNSGYLDNDNCARITLIFNGKNVPQ	959		
Qy	973	GLATTEHICSELRNLNLIKNSABYSYIACEPSPSANNEIHVAISABDIRD--DGNPIKE	1030		
Db	960	GTTTESICSEIRYLPAFESVSRDRTLILCDLSYSTENAVEAISFPVHRDEQDNLQION	1019		
Qy	1031	ITDKIIDLVSQRDGNSSLIAAEVRRVQ	1058		
Db	1020	AANITVNAVTKRO-NSTWMLAVTEVKVE	1046		

```

RESULT 12
Q90Y56 BRARE PRELIMINARY; PRT; 1254 AA.
ID Q90Y56 BRARE PRELIMINARY; PRT; 1254 AA.
AC Q90Y56;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Jagged2.
GN Name=Jag2;
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
NCBI_taxid=7955;
[1]
RN NUCLEOTIDE SEQUENCE.
RA Oda T., Chandrasekharappa S.C.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
EMBL: AF729449; AAL08214.1; -, mRNA.
DR HSP; P01132; IGK5.
DR Ensembl; ENSDARG0000021389; Danio rerio.
DR ZFIN; ZDB-GENE-011128-3; jag2.
GO; GO:0016020; C:membrane; IEA.
GO; GO:0005509; F:calcium ion binding; IEA.
GO; GO:0007154; P:cell communication; IEA.
DR InterPro; IPR000152; Aa_hydroxyl_S.
DR InterPro; IPR001774; DSL.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR001881; EGF CA.
DR InterPro; IPR001438; EGF-II.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR011651; MNFL.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF01414; DSL; 1.
DR Pfam; PF00008; EGF; 14.
DR Pfam; PF07645; EGF CA; 1.
DR Pfam; PF07657; MNFL; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR00010; EGFLOOD.
DR SMART; SM00051; DSL; 1.
DR SMART; SM00179; EGF CA; 8.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 7.
DR PROSITE; PS00022; EGF_1; 16.
DR PROSITE; PS01186; EGF_2; 11.
DR PROSITE; PS50026; EGF_3; 15.
DR PROSITE; PS01187; EGF CA; 6.
DR PROSEQUENCE 1254 AA; 137719 MW; AFF70717DF190FEB CRC64;

```

Query Match	60.7%	Score	3795.5	DB 2	Length	1254			
Best Local Similarity	58.0%	Pred. No.	4.3e-249						
Matches	622	Conservative	171	Mismatches	255	Indels	25	Gaps	11
Qy	7	RGRSGRPLSLILLALLCALRAKVCASQCFLEILLSMNVNGELQNGNCCGARGNPGDKC	66						
Db	6	RIRNLWLPFA---CLLLTMTWKVSQSSGYFELQIAVENVNGELWDGECDSSTRNSQDQRC	62						
Qy	67	TRRECDTYFKVCLKEYQSRVYTAGVPCPSFGSGSTPVIIGNTFNLKA-----SRGNDNRNIV	121						
Db	63	VRRECDTYFKVCLKEYQSEVTTTGGCTFFGSGSTDVLGNIFFSFKTAKNSPKTSDVGKII	122						
Qy	122	LPFSFAWPRSYTLLVRAWDSSNDTVQ--PDSIITEKASHSGMINPSRQWOTLKQNTGVVAHF	179						
Db	123	IPFHFAWPRSYTLLILEAWDNDNSNTQNGEENLIERHIASHMVNPGDHWQIHPFGITAH	182						
Qy	180	EYQIRVTCDDYYVGFGCNKFCRRPDDDPFGHVACDQNGNKKTCMEGWMPCECNRAJCRQSCS	239						
Db	183	EYRIRVREDENYVYGSKNCKQCRPRDDYFGHYRCDPSGNI VCLDGMWGEDCKRTALCKQGCN	242						
Qy	240	PKHGSKLPGDRCQYWGQGLYCDKCIPIHPFGCVHGI CNBPWQCLCETNWWGSQLCDKQINY	299						


```
Db 243 LIHGCAVPGECKCNVYQWQGFCDCECLPYFGLHGTCTVMPWQCTCEKQWGLLCKDLNY 302
Qy 300 CGTHQPCNLNGTSCNTCPDKYQSCPEGYSGPNCIEAHACLSDPCHNRGSCKETSLGFE 359
Db 303 CGTHPCVNGGTCTMNSPEDEYNACPEGYSGKNCIEAHACVSNPCANGGTCHVPTGFE 362
Qy 360 CEGSPGWGTCTSTNIDDCSPNNCSHGCTQDILVNGFKVCPQWGTCTKOLDANECEAK 419
Db 363 CHCPGWEGFTCAKMDDECASSPCAQGTCTDILENGFECVCPQWVGKTCQIDANECEAK 422
Qy 420 PCVNAKSKNLIIASYYCDCLPGWNGQNCININDLQCGQNDASCRDLVN-GYRCICPPG 478
Db 423 PCVNAHCKNMIIGYHCDCFQWAGQNCIDINLNGCHGQCGATCKELVHGYYHCQCPAG 482
Qy 479 YAGDHCDERDIDECASNPLNGCHGCHQNEINRFQCLPTGFGNLCQLDI----DYCEPNPC 534
Db 483 FVGLHCEVSRNKACSGPCQNGRCHVLDLDFVCECPNSYAGMLCEVESLHPNCEPNPC 542
Qy 535 QNGAQCYNRASDYFCKPEDEYEGKNCSHLKHCRKTTTCEVIDSTCTVAMASNDTPEGVRYI 594
Db 543 QNTALCYSLPGDFYFACPEDEYEGTCENRKHCRKTTTCEVIDSTCTVAMASNDTPEGVRYI 602
Qy 595 SSNVCGPHGCKKSQSGKFTCDCKMGFTGTGTCHENINDCESNPCRNGGTCTDGVNSYKCI 654
Db 603 NSNVCGPHGRCISQPGNFTCTCELGFTGTGTCHENVDCVSNPCRNGGTCTDGVNSYKCI 662
Qy 655 CSQDWEAGYCEETNINDCSQNPCHNGTCRDLVNDFYCDCKNGKWKGTCHSRDQCEATC 714
Db 663 CPDWEGLDLSINVNECSRSPCKNGRCHVLDLNDFYCECANGKWKGTCHSRDQCEATC 722
Qy 715 NNGGTCYDEGDARCKMCPGWEGTTCNARNSSCLPNPCHNGGTCTVNGSGFTCVCKEGW 774
Db 723 SNGGTCYDHGDAFRCACPPGEGSTCTNANKNSCASGCLNGGTCTVNGSGFTCTVCKEGW 782
Qy 775 EGPICATNTNDCSPPHPCVNSGTCVDGNWYRCECAPGAGFAGPDCRININEQSSPCAFGAT 834
Db 783 EGATCAQNTNDCNPHPCVNGGICVDGNWYRCECAPGAGFAGPDCRININEQSSPCAFGAT 842
Qy 835 CVDBINGYRCVCPGHGSAKQEV--SGRPCITMGSVIPDGAQWDDCNTCQNLNGRIAC 892
Db 843 CVDBINGYRCVCPGLRGTGQPEIGIGTKHYAGLQFPHGSRWEEECNTCCQVNGKVE 902
Qy 893 SKWCGPRPCLL---HKGHSECPGSGOSCPILDQCFVHPCTGVGBCRSSLQOP-VKTK 947
Db 903 TKVVGCRKPCLLPOTPGRELHYCGGREGLEHFLTCLSPCHQWGFCSPEPTPTQTK 962
Qy 948 CTSDSYQD--NCANITFTFKEMMSPLTTEHICSELRLNMLKNVSAEYSIYIACBPSP 1006
Db 963 CEPNSVYLDKSCARITLIFNRDLKPTGTTVENVCSELRYMPATRSIAKDHKLLVLCDSY 1022
Qy 1007 SANNEIHAISAE--DIRDDGNPIKEITDKIIDLVSKRDNSSLLIAAANVAVRQ 1058
Db 1023 KNODAVEVAISFOHDEQPSHSQIOEAASITVSTLSKRH--NSTVMIAVIEVKVE 1074

RESULT 13
Q5TZK8_BRARE PRELIMINARY; PRT; 1254 AA.
ID Q5TZK8 BRARE PRELIMINARY;
AC Q5TZK8;
DT 01-FEB-2005 (TReMBLrel. 29, Created)
DT 01-FEB-2005 (TReMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TReMBLrel. 29, Last annotation update)
DE Jagged2.
GN Name=Jag2; ORFNames=DKXY-5P1.1-001;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_taxid=7955;
RN NUCLEOTIDE SEQUENCE.
RA Corby N.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; BX004766; CAH69087.1; -; Genomic_DNA.
DR GO; GO:0016020; C-membrane; IEA.
DR GO; GO:000509; F-calciun ion binding; IEA.
DR GO; GO:0007154; P-cell communication; IEA.
DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR001174; DSL.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR011651; MNHL.
DR InterPro; IPR006552; VWC_out.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF01414; DSL; 1.
DR Pfam; PF00008; EGF; 14.
DR Pfam; PF07645; EGF_CA; 1.
DR Pfam; PF07657; MNHL; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR00010; EGFLOOD.
DR SMART; SM00051; DSL; 1.
DR SMART; SM00181; EGF; 16.
DR SMART; SM00179; EGF_CA; 14.
DR SMART; SM00215; VWC_out; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 7.
DR PROSITE; PS00022; EGF_1; 16.
DR PROSITE; PS01186; EGF_2; 11.
DR PROSITE; PS00026; EGF_3; 15.
DR PROSITE; PS01187; EGF_CA; 6.
DR PROSITE; PS01187; EGF_CA; 6.
SQ SEQUENCE 1254 AA; 137623 MW; D546358P912C9FFA CRC64;

Query Match 60.7%; Score 3790.5; DB 2; Length 1254;
Best Local Similarity 57.9%; Pred. No. 9.3e-249;
Matches 621; Conservative 171; Mismatches 256; Indels 25; Gaps 11;

Qy 7 RGRSGPLSLALLCALRAKVCASGQFELEILSMQNVGELQNGNCCGARNPGDRKC 66
Db 6 RIRNLPPIA---CLLLTMTWTKVQSQSGYFELQLIAENVNGLWDGECDDSTRNSQDRC 62
Qy 67 TRDECDTYFKVCLKEYQSRVTAAGPSCFGSGSPVIGGNTFNLA-----SRGNDNRIV 121
Db 63 VRDECDTYFKVCLKEYQSEVTTTGGCTFGSGSTDVLGNIISFKTAKNSPKTSDVGGKII 122
Qy 122 LPFSFAWPRSYTLLEVAWDSNDTVQ--PDSIIKASHSGMINPSPROWTLKQNTGVAFH 179
Db 123 IPFHFAWPRSYTLLEVAWDSNDTVQ--PDSIIKASHSGMINPSPROWTLKQNTGVAFH 182
Qy 180 EYQIRVTCDYFYFGCNKFCRPRDDPFPHYACDQNGKTCMBGWMGPECNRAICRGCS 239
Db 183 EYRIRVRCDENYVYSGKNCQCRPRDDYFHYRCDPSGNIIVCLDGMWGEDCRTAICKQCN 242
Qy 240 PKHGSCKLPDRCQYQWQGLYCDKCLIPHPGCVHIGICNEPWQCLCETNWGGQLCDKLN 299
Db 243 LIHGCAVPGECKCNVYQWQGFCDCECLPYFGLHGTCTVMPWQCTCEKQWGLLCKDLNY 302
Qy 300 CGTHQPCNLNGTSCNTCPDKYQSCPEGYSGPNCIEAHACLSDPCHNRGSCKETSLGFE 359
Db 303 CGTHPCVNGGTCTMNSPEDEYNACPEGYSGKNCIEAHACVSNPCANGGTCHVPTGFE 362
Qy 360 CEGSPGWGTCTSTNIDDCSPNNCSHGCTQDILVNGFKVCPQWGTCTKOLDANECEAK 419
Db 363 CHCPGWEGFTCAKMDDECASSPCAQGTCTDILENGFECVCPQWVGKTCQIDANECEAK 422
Qy 420 PCVNAKSKNLIIASYYCDCLPGWNGQNCININDLQCGQNDASCRDLVN-GYRCICPPG 478
Db 423 PCVNAHCKNMIIGYHCDCFQWAGQNCIDINLNGCHGQCGATCKELVHGYYHCQCPAG 482
Qy 479 YAGDHCDERDIDECASNPLNGCHGCHQNEINRFQCLPTGFGNLCQLDI----DYCEPNPC 534
Db 483 FVGLHCEVSRNKACSGPCQNGRCHVLDLDFVCECPNSYAGMLCEVESLHPNCEPNPC 542
Qy 535 QNGAQCYNRASDYFCKPEDEYEGKNCSHLKHCRKTTTCEVIDSTCTVAMASNDTPEGVRYI 594
```



```
QY 715 NNGTCYDEGDAKCMPCGGMEGTCNTAIENSSCLPNPCHNGGTCVYNGESFTCVCKEGW 774
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
723 NNGGTCYDHDGAFRCACPPGWEGSTCTAKNSTCASGCLNGGTCVGGGDTFTCIKRDGW 782
QY 775 EGPICANTNDSCPHPCYNCGTICVDGNWYRCECAGFAGPDCRININEQSSPCAFGAT 834
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
783 EGATCAQNTDCHPHPCYNGGICVDGNWYRCECAGFAGPDCRININEQSSPCAYGAT 842
QY 835 CVDEINGYRCVCPGHSAGKQEV--SGRPCITWGSVIPDGAKWDDCNTCQCLNGRIAC 892
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
843 CVDEINGYRCVCPGLRTPGQCFEIGIKTCHVAGLQFPHGSWEBCNCTCQCVNGKVEC 902
QY 893 SKWCGRRPCLL-----HKHSECPGSGOSCPILDDQCFVHPCTGVGECRSSSLQP-VKTK 947
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
903 TKVCGRGPCLLPOTGPORELHYCPGGRECHFNFLTCLSPCHQWGPCSSPETPTLQTK 962
QY 948 CTDSYYQD--NCANITFTFNKEMWSPGLTTEHCSELNRLNKLWNSAEYSIYIACEPSP 1006
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
963 CEPNSVLDKSCARITILFNKDLPTGTTVENVCSELRYMPATRSKADKHLVLCDLSY 1022
QY 1007 SANNEIHVAISAE--DIRDDGNPKIEITDKIIDLVSKRDNSSLIAAVAEVRVQ 1058
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1023 KNQDAVEVAISFQHQEQSHSQIEAAASTIVSTLSKRH-NSTVMLAVIEVKVE 1074

RESULT 15
JAG2 HUMAN
ID --JAG2 HUMAN STANDARD; PRT; 1238 AA.
AC Q9Y2I9; Q9UE17; Q9UE99; Q9UNK8; Q9Y6P9; Q9Y6Q0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Jagged-2 precursor (Jagged2) (HJ2).
GN Name=JAG2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
(1)
NCLEOTIDE SEQUENCE (ISOFORM LONG).
MEDLINE=97459705; PubMed=9315665;
Luo B., Aster J.C., Haasarian R.P., Kuo F., Sklar J.;
"Isolation and functional analysis of a cDNA for human Jagged2, a gene
encoding a ligand for the Notch1 receptor.";
Mol. Cell. Biol. 17:6057-6067(1997).
(2)
NCLEOTIDE SEQUENCE (ISOFORM LONG).
TISSUE=Petal brain;
MEDLINE=99180765; PubMed=10079256;
Gray G.E., Mann R.S., Mitsiadis E., Henrique D., Carcangiu M.-L.,
Banks A., Leiman J., Ward D., Ish-Horowitz D., Artavanis-Teakonas S.;
"Human ligands of the Notch receptor.";
Am. J. Pathol. 154:785-794(1999).
(3)
NCLEOTIDE SEQUENCE (ISOFORMS LONG AND SHORT).
TISSUE=Bone marrow;
MEDLINE=20130121; PubMed=10662552; DOI=10.1006/geno.1999.6045;
Deng Y., Madan A., Banta A.B., Friedman C., Trask B.J., Hood L.,
Li L.;
"Characterization, chromosomal localization, and the complete 30-kb
DNA sequence of the human Jagged2 (JAG2) gene.";
Genomics 63:133-138(2000).
(4)
NCLEOTIDE SEQUENCE OF 17-1238 (ISOFORM LONG).
TISSUE=Heart;
MEDLINE=98145947; PubMed=9486542; DOI=10.1016/S0925-4773(97)00146-9;
Valsecchi C., Ghezzi C., Ballabio A., Rugari E.I.;
"JAGGED2: a putative Notch ligand expressed in the apical ectodermal
ridge and in sites of epithelial-mesenchymal interactions.";
Mech. Dev. 69:203-207(1997).
CC -!- FUNCTION: Putative Notch ligand involved in the mediation of Notch
```

signaling. Involved in limb development (By similarity).

-!- SUBCELLULAR LOCATION: Type I membrane protein.

-!- ALTERNATIVE PRODUCTS:

Event=Alternative splicing; Named isoforms=2;

Name=Long;

isoId=Q9Y2I9-1; Sequence=Displayed;

Name=Short; Synonyms=HJAG2.del-E6;

isoId=Q9Y2I9-2; Sequence=VSP 001395;

-!- TISSUE SPECIFICITY: Expressed in heart, placenta and skeletal muscle and to a lesser extent in pancreas. Very low expression in brain, lung, liver and kidney.

-!- DISEASE: May be associated with Usher syndrome type IA (USH1A1A) which describes a congenital sensory deafness associated with retinitis pigmentosa and feeble-mindedness.

-!- SIMILARITY: Contains 1 DSL domain.

-!- SIMILARITY: Contains 16 EGF-like domains.

-!- SIMILARITY: Contains 1 WMFC domain.

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

EMBL; AF020201; AAB71189.1; -; mRNA.

EMBL; AF003521; AAB61285.1; -; mRNA.

EMBL; AF029778; AAB84215.1; -; mRNA.

EMBL; AF029779; AAB84216.1; -; mRNA.

EMBL; AF111170; AAD15562.1; -; Genomic_DNA.

EMBL; Y14330; CAA74706.1; -; mRNA.

HSSP; P01132; 1GK5.

Ensembl; ENSG00000184916; Homo sapiens.

HGNC; HGNC:5189; JAG2.

Reactome; Q9Y2I9; -.

MIT; 602570; -.

GO; GO:0005887; C:integral to plasma membrane; ISS.

GO; GO:0008083; F:growth factor activity; IDA.

GO; GO:0005112; F:Notch binding; IPI.

GO; GO:000912; P:auditory receptor cell fate commitment; ISS.

GO; GO:0007049; P:cell cycle; NAS.

GO; GO:0030154; P:cell fate determination; IDA.

GO; GO:0001709; P:cell fate determination; NAS.

GO; GO:0007267; P:cell-cell signaling; ISS.

GO; GO:0003026; P:embryonic limb morphogenesis; ISS.

GO; GO:0008583; P:mystery cell fate differentiation (sensu En. .; TAS.

GO; GO:0007219; P:Notch signaling pathway; NAS.

GO; GO:0007605; P:perception of sound; ISS.

GO; GO:0030334; P:regulation of cell migration; NAS.

GO; GO:0042127; P:regulation of cell proliferation; IDA.

GO; GO:0007283; P:spermatogenesis; IEP.

GO; GO:0030217; P:T cell differentiation; IDA.

GO; GO:0045861; P:thymic T cell selection; IDA.

InterPro; IPR000152; Asx_hydroxyl_S.

InterPro; IPR001774; DSL.

InterPro; IPR001881; EGF_Ca_bd.

InterPro; IPR001438; EGF_II.

InterPro; IPR006209; EGF_like.

InterPro; IPR011651; MNFL.

Pfam; PF01414; DSL; 1.

Pfam; PF00008; EGF; 12.

Pfam; PF07645; EGF_CA; 1.

Pfam; PF07657; MNFL; 1.

PRINTS; PR00010; EGFBLD.

PROSITE; PS00010; ASX_HYDROXYL; 10.

PROSITE; PS1051; DSL; 1.

PROSITE; PS00022; EGF_1; 16.

PROSITE; PS01185; EGF_2; 12.

PROSITE; PS00026; EGF_3; 15.

PROSITE; PS01187; EGF_CA; 7.

PROSITE; PS01208; WMFC_1; FALSE_NEG.

PROSITE; PS0184; WMFC_2; 1.

KW Alternative splicing; Calcium; Developmental protein; EGF-like domain;
KW Glycoprotein; Notch signaling pathway; Repeat; Signal; Transmembrane.
FT SIGNAL 1 26 Potential.
FT CHAIN 27 1238 Jagged-2.
FT TOPO_DOM 27 1080 Extracellular (Potential).
FT TRANSMEM 1081 1105 Potential.
FT TOPO_DOM 1106 1238 Cytoplasmic (Potential).
FT DONAIN 196 240 DSL.
FT DONAIN 241 274 EGF-like 1.
FT DONAIN 275 305 EGF-like 2.
FT DONAIN 307 345 EGF-like 3.
FT DONAIN 347 383 EGF-like 4.
FT DONAIN 385 421 EGF-like 5; calcium-binding (Potential).
FT DONAIN 423 459 EGF-like 6; calcium-binding (Potential).
FT DONAIN 461 496 EGF-like 7; calcium-binding (Potential).
FT DONAIN 498 534 EGF-like 8.
FT DONAIN 536 572 EGF-like 9.
FT DONAIN 574 634 EGF-like 10; atypical.
FT DONAIN 636 672 EGF-like 11; calcium-binding (Potential).
FT DONAIN 674 710 EGF-like 12; calcium-binding (Potential).
FT DONAIN 712 748 EGF-like 13.
FT DONAIN 751 787 EGF-like 14.
FT DONAIN 789 825 EGF-like 15; calcium-binding (Potential).
FT DONAIN 827 863 EGF-like 16; calcium-binding (Potential).
FT DONAIN 870 944 WFC.
FT CARBOHYD 153 153 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 570 570 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 619 619 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 752 752 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 1058 1058 N-linked (GlcNAc. . .) (Potential).
FT DISULFID 245 256 By similarity.
FT DISULFID 249 262 By similarity.
FT DISULFID 264 273 By similarity.
FT DISULFID 276 287 By similarity.
FT DISULFID 282 293 By similarity.
FT DISULFID 295 304 By similarity.
FT DISULFID 311 323 By similarity.
FT DISULFID 317 333 By similarity.
FT DISULFID 335 344 By similarity.
FT DISULFID 351 362 By similarity.
FT DISULFID 356 371 By similarity.
FT DISULFID 373 382 By similarity.
FT DISULFID 389 400 By similarity.
FT DISULFID 394 409 By similarity.
FT DISULFID 411 420 By similarity.
FT DISULFID 427 438 By similarity.
FT DISULFID 432 447 By similarity.
FT DISULFID 449 458 By similarity.
FT DISULFID 465 475 By similarity.
FT DISULFID 469 484 By similarity.
FT DISULFID 486 495 By similarity.
FT DISULFID 502 513 By similarity.
FT DISULFID 507 522 By similarity.
FT DISULFID 524 533 By similarity.
FT DISULFID 540 551 By similarity.
FT DISULFID 545 560 By similarity.
FT DISULFID 562 571 By similarity.
FT DISULFID 589 612 Potential.
FT DISULFID 606 622 Potential.
FT DISULFID 624 633 By similarity.
FT DISULFID 640 651 By similarity.
FT DISULFID 645 660 By similarity.
FT DISULFID 662 671 By similarity.
FT DISULFID 678 689 By similarity.
FT DISULFID 683 698 By similarity.
FT DISULFID 700 709 By similarity.
FT DISULFID 716 727 By similarity.
FT DISULFID 721 736 By similarity.
FT DISULFID 738 747 By similarity.
FT DISULFID 755 766 By similarity.
FT DISULFID 760 775 By similarity.
FT DISULFID 777 786 By similarity.
FT DISULFID 793 804 By similarity.

FT DISULFID 798 813 By similarity.
FT DISULFID 815 824 By similarity.
FT DISULFID 831 842 By similarity.
FT DISULFID 851 861 By similarity.
FT DISULFID 853 862 By similarity.
FT VARSPLIC 424 461 Missing (in isoform Short).
FT CONFLICT 8 12 /FTID=VSP_001395.
FT CONFLICT 119 119 RLPRR -> APPPA (in Ref. 1).
FT CONFLICT 119 119 A -> P (in Ref. 1).
Query Match 59.0%; Score 3685; DB 1; Length 1238;
Best Local Similarity 57.3%; Pred. No. 1.3e-241;
Matches 619; Conservative 153; Mismatches 272; Indels 36; Gaps 13;
QY 5 RTRSGRPLSLALLCALRAKVCASGQFELEILSMQNVGELQNGCC-GGARNPGD 63
DB 2 RAQGRGLPRLL--LJLLALWQAARPMGYFELQLSALRVNVELLSGACDGDGRTRA 59
QY 64 RKTRDECTYFKVCLKEYQSRVTAGGPCSGFGSGSTPVTCGNTFNL-----KA 111
DB 60 GCGCHDECDTYVRVCLKEYQAKVTPGTGSCYGHGATPVLGNSFYLPAGAGDRARARA 119
QY 112 SRGNDRN--RIVLPFSFAMPRSYTLVAEAWDSNDTV-QPDSIIERKASHGMINPSQWQ 168
DB 120 RAGGDQDPLGLVLPQFAMPRSFLLIVEAWDNDTTPNEELLIERVSHAGMINPDRWK 179
QY 169 TLKQNTGVAHFQIRVTCDYVYFGCNKFCRPRDDFFGHYACDQNGNKTCEGMWGP 228
DB 180 SLHFSGHVALELQIRVCDENYYSATCNKFCRPRDDFFGHYTCQYGNKACMDGWNGKE 239
QY 229 CNRAICRQGCSPKHGSKLPGDCRCQYWGQLYCDKCIPIHPGCVHGINPWCCLCTNW 288
DB 240 CREAVCKQGNLLHGGCTVPGECRCSYGWQRCPCDECVPGCVHSGSCVPEWQNCSTNW 299
QY 289 GGQCDKDLNYCGTHOPCLNGGTCNTGDKVQCSCEGYSNGNCEIAEACILSDPCNR 348
DB 300 GGLLCDKDLNYCGSHHPCTNGGTCINAEPPQVRCCTPDGYSGRNCKAEHACTSNPCANG 359
QY 349 GSCKETSLGFECECPGWTGPTCSTNIDDCSPNNCSHGCTCQDLVNGFKVCVCPQWTGKT 408
DB 360 GSCEVPSGFECHCPGSGMGPCTALDIDECASNPAAAGTCVQDVDFEICPEQWVGAT 419
QY 409 QQLDANECAKPCVNAKSKNLIASYYCDCLPGWQNGQCDININDCLGQCNDSASCDLV 468
DB 420 QQLDANECEGKPCLNAPFCKNLIGYYCDICPGWKGINCHINVDNCRGQCQHGCTCKLV 479
QY 469 NGYRCICPGYAGDHCEERDIDECASNPCLNGHCQNEINRQCICPTGSEGNLCQDIDY 528
DB 480 NGYQCVCPRGFRGHRHCELRDKCASPCHSGGLCDLADGFHCHCPQGFGLCEVDVLD 539
QY 529 CEPNPNQGAQCVNRASDYFCKCPEDYEGKNCSHLKHCRTRTPCEVIDSCVTAMASNDTP 588
DB 540 CFPSPCENGARCVNLEGDYCYACPDPDFGKNCVSPREPCEGACRVIDGC----GSDAGP 595
QY 589 EGVYIISNVCGPHGKCKSQSGKFTCDCKNGTGTGYCHENINDCESNPNCRNGGTCDGV 648
DB 596 GMPGTAASGVCGRVGRVSPQGNFSCIQDSGTGTGYCHENIDCLGQPCRNNGGTCDIEV 655
QY 649 NSYKCI CSQGWEGAYCETINIDCSQNPCHNGGTCRDLVNDYCDCKNGKWKGTCHSRDSQ 708
DB 656 DAFCRCFPGSGEGLCDTNPNDCLPDPCHSRGRCYDLVNDYFACDDGKWKGTCHSRFQ 715
QY 709 CDEATCNNNGTVCYDEGDAFKCMCPGWEGTTCNIARNSSCLPNPCHNGGTCVVVNGSFTC 768
DB 716 CDAYTCNNGT CYDSGDTFRACPPGWKGTSTAVAKNSSCLPNPCVNGGTCVSGSASFSC 775
QY 769 VCKEGWEGPICAQNTNDCSPHPCYNSGTCTVDGNWYRCCECAPGAGPCRCRININEQSSP 828
DB 776 ICRDGEWGRCTHTNTDNCNPLPCYNGGICVDGWNWFRCECAPGAGPCRCRINIDECQSSP 835
QY 829 CAFGATCVDEINGYRCVCPPHSGAKCOEV--SGRBCITMGSVIPDGAKWDDCNCQCL 886
DB 836 CAYGATCVDEINGYRCSCPGRAGRCQEVIGFRSCWSRGTPFPHGSSWVEDCNCRCCL 895

Qy	887	NGRIACSKVWCGPRPCLLLHKGHSE-----CPSGOSCIPIILDDQCFVHPCTGVGECRSSSL	941
Db	896	DGRRDCSKVWCGWKPCLL-AGQPEALSAQCPLGQRCLEKAPGQCLRPPCEAWGECGAE--	952
Qy	942	QPVKTKTSDSYOD-NCANITFTENKEMMSPLTTEHICSELNRLNLIKNVSAEYSIVI	1000
Db	953	EPPSTPCLPRSGHLDNNCARLTLHFNRDHVPQGTTVGALCSGIRSLPATRAVARDRLVL	1012
Qy	1001	ACEPSPSANNEIHVAISAEDIRD--DGNPIKEITDKIIDLVSKRDGNSSLIAAAVAEVRVQ	1058
Db	1013	LCDRASSGASAVEVAVSFSPARDLPDSSLIQGAHAIVAIAITQR-GNSSLLLAATEVKVE	1071

Search completed: February 4, 2006, 09:35:31
Job time : 276 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2006, 09:06:39 ; Search time 645 Seconds
(without alignments)

10078.361 Million cell updates/sec

Title: US-10-650-650-2

Perfect score: 3655.8

Sequence: 1 atgcgttccacggacrcg.....gaatggagtagcatcgtagatg 3657

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing files: 45 summaries

Database : Issued Patents NA*

- 1: /cgn2_6/prodata/1/ina/1_COMB.seq.*
- 2: /cgn2_6/prodata/1/ina/5_COMB.seq.*
- 3: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/prodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/prodata/1/ina/6H_COMB.seq.*
- 6: /cgn2_6/prodata/1/ina/PCRTUS_COMB.seq.*
- 7: /cgn2_6/prodata/1/ina/PP_COMB.seq.*
- 8: /cgn2_6/prodata/1/ina/RE_COMB.seq.*
- 9: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3655.8	100.0	3657	US-09-579-536C-2	Sequence 2, Appli
2	3654.2	99.9	5590	US-08-882-046-1	Sequence 1, Appli
3	3654.2	99.9	5590	US-09-566-047-1	Sequence 1, Appli
4	3654.2	99.9	5896	US-09-949-016-31	Sequence 31, Appli
5	3651	99.9	4208	US-09-068-740A-10	Sequence 10, Appli
6	3651	99.9	5885	US-09-949-016-4426	Sequence 4426, Ap
7	3649.4	99.8	4855	US-09-917-254-34	Sequence 34, Appli
8	3647.8	99.8	4208	US-09-214-278-6	Sequence 6, Appli
9	3647.8	99.8	4208	US-09-855-722-6	Sequence 6, Appli
10	3646.2	99.7	6464	US-08-400-159-5	Sequence 5, Appli
11	3646.2	99.7	6464	US-08-611-729A-5	Sequence 5, Appli
12	3646.2	99.7	6464	US-09-195-524-5	Sequence 5, Appli
13	3646.2	99.7	6464	US-09-310-685-3	Sequence 3, Appli
14	3611.8	98.8	5458	US-09-199-865-2	Sequence 2, Appli
15	3611.8	98.8	5458	US-10-213-329-2	Sequence 2, Appli
16	3199.8	87.5	3201	US-09-579-536C-17	Sequence 17, Appli
17	2455.4	67.2	3582	US-08-400-159-9	Sequence 9, Appli
18	2455.4	67.2	3582	US-08-611-729A-9	Sequence 9, Appli
19	2455.4	67.2	3582	US-09-195-524-9	Sequence 9, Appli
20	2455.4	67.2	3582	US-09-310-685-7	Sequence 7, Appli
21	1157.8	31.7	3955	US-09-214-278-4	Sequence 4, Appli
22	1157.8	31.7	3955	US-09-855-722-4	Sequence 4, Appli
23	1056.4	28.9	4483	US-08-611-729A-7	Sequence 7, Appli
24	1056.4	28.9	4483	US-09-195-524-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1

US-09-579-536C-2

Sequence 2, Application US/09579536C

Patent No. 6716974

GENERAL INFORMATION:

APPLICANT: MACIAG, Thomas

APPLICANT: ZIMRIN, Ann

APPLICANT: SMALL, Deena

APPLICANT: PRUDOVSKY, Igor

TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC METHODS AND COMPOSITIONS BASED ON JAGT

TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS

FILE REFERENCE: 053689-5002-01

CURRENT APPLICATION NUMBER: US/09/579,536C

CURRENT FILING DATE: 2000-05-24

PRIOR APPLICATION NUMBER: US 09/199,865

PRIOR FILING DATE: 1998-11-25

PRIOR APPLICATION NUMBER: PCT/US97/09407

PRIOR FILING DATE: 1997-05-30

PRIOR APPLICATION NUMBER: US 60/018,841

PRIOR FILING DATE: 1996-05-31

NUMBER OF SEQ ID NOS: 56

SOFTWARE: PatentIn version 3.1

SEQ ID NO 2

LENGTH: 3657

TYPE: DNA

ORGANISM: Homo sapiens

US-09-579-536C-2

Query Match 100.0%; Score 3655.8; DB 3; Length 3657;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3657; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGCGTTCACCGACRCGCGCGGTCGCGGCGCCCTAGCCCTGCTGCTGCGCCCTG 60
DB	1	ATGCGTTCACCGACRCGCGCGGTCGCGGCGCCCTAGCCCTGCTGCTGCGCCCTG 60
QY	61	CTCTGTGCGCCTGCGAGCCAAAGTGTGTGTGGGGCTCGGGGTGAGTTCGAGATCCTG 120
DB	61	CTCTGTGCGCCTGCGAGCCAAAGTGTGTGTGGGGCTCGGGGTGAGTTCGAGATCCTG 120
QY	121	TCATGACAGACGTGAACCGGAGAGTGCAGAACCGGAACTGTGCGCGCGCGCCGGAAC 180
DB	121	TCATGACAGACGTGAACCGGAGAGTGCAGAACCGGAACTGTGCGCGCGCGCCGGAAC 180
QY	181	CGGGAGACCGCAAGTGCACCCCGGACGAGTGTGACACATCTTCAAGTGTGCTCAAG 240
DB	181	CGGGAGACCGCAAGTGCACCCCGGACGAGTGTGACACATCTTCAAGTGTGCTCAAG 240

QY	241	GAGTATCAGTCCGCGTCAAGCGCGGGGCGCTTGCAGCTTCGGCTCAGGGTCCAGCCT	300
Db	241	GAGTATCAGTCCGCGTCAAGCGCGGGGCGCTTGCAGCTTCGGCTCAGGGTCCAGCCT	300
QY	301	GTCAATCGGGGCAACACCTTCAACCTCAAGCCAGCGCGGCAACGACCGCAACCGCATC	360
Db	301	GTCAATCGGGGCAACACCTTCAACCTCAAGCCAGCGCGGCAACGACCGCAACCGCATC	360
QY	361	GTGCTGCTTTTCAGTTTGGCTGGCGAGGTCTCTATACGTGTTGCTTGTGAGGGCTGGGAT	420
Db	361	GTGCTGCTTTTCAGTTTGGCTGGCGAGGTCTCTATACGTGTTGCTTGTGAGGGCTGGGAT	420
QY	421	TCCAGTAATGACACCGTTCAACCTGACAGTATTTATGAAAGGCTTCTCACTCGGCGATG	480
Db	421	TCCAGTAATGACACCGTTCAACCTGACAGTATTTATGAAAGGCTTCTCACTCGGCGATG	480
QY	481	ATCAACCCAGCGCGAGTGGCAGACGCTCAAGCAGAAACAGCGGCGTTCGCCATTTGAG	540
Db	481	ATCAACCCAGCGCGAGTGGCAGACGCTCAAGCAGAAACAGCGGCGTTCGCCATTTGAG	540
QY	541	TATCAGATCCGCGTACCTGTGATGACTACTATGCGCTTTGGCTGTAATAAGTCTGTC	600
Db	541	TATCAGATCCGCGTACCTGTGATGACTACTATGCGCTTTGGCTGTAATAAGTCTGTC	600
QY	601	CGCCCGAGATGACTTCTTTGACACTATGCGCTGTGACAGAAATGGCAAAACTTGC	660
Db	601	CGCCCGAGATGACTTCTTTGACACTATGCGCTGTGACAGAAATGGCAAAACTTGC	660
QY	661	ATGGAAGGCTGGATGGGCCCGGATTAACAGAGCTATTTGCCGACAAAGGCTCAGTCTCT	720
Db	661	ATGGAAGGCTGGATGGGCCCGGATTAACAGAGCTATTTGCCGACAAAGGCTCAGTCTCT	720
QY	721	AAGCATGGGTCTTGCAAACTCCAGGTGACTGCAAGTGCAGTAYGGCTGGCAAGGCTG	780
Db	721	AAGCATGGGTCTTGCAAACTCCAGGTGACTGCAAGTGCAGTAYGGCTGGCAAGGCTG	780
QY	781	TACTGTGATAGTGCATCCACACCGGGATGGTCCACGGCATCTGTATAGACCTGG	840
Db	781	TACTGTGATAGTGCATCCACACCGGGATGGTCCACGGCATCTGTATAGACCTGG	840
QY	841	CAGTGCCTCTGTGAGACCACTGGGCGCGCAGCTCTGTGACAAAGATCTCAATTAAGT	900
Db	841	CAGTGCCTCTGTGAGACCACTGGGCGCGCAGCTCTGTGACAAAGATCTCAATTAAGT	900
QY	901	GGGACTCATAGCGGTGTCTCAACGGGGGAACTTTGTAGCAACACAGGCGCTGCAAAAT	960
Db	901	GGGACTCATAGCGGTGTCTCAACGGGGGAACTTTGTAGCAACACAGGCGCTGCAAAAT	960
QY	961	CAGTGTTCCTGCGCTGAGGGGTATTCAGGACCCAACTGTGAAATTTGTGAGCAGCGCTGC	1020
Db	961	CAGTGTTCCTGCGCTGAGGGGTATTCAGGACCCAACTGTGAAATTTGTGAGCAGCGCTGC	1020
QY	1021	CTCTCTGATCCCTGTCAACAGAGGAGCTGTAAAGAGACCTCCCTGGGCTTTGAGTGT	1080
Db	1021	CTCTCTGATCCCTGTCAACAGAGGAGCTGTAAAGAGACCTCCCTGGGCTTTGAGTGT	1080
QY	1081	GAGTGTTCCTGAGCTGGAACGGGCCCAATGCTCTCAAAATTTGATGACTGTTCTCTCT	1140
Db	1081	GAGTGTTCCTGAGCTGGAACGGGCCCAATGCTCTCAAAATTTGATGACTGTTCTCTCT	1140
QY	1141	AATAACTGTTCCACAGGGGCACTGCGAGGACCTGGTTAACGGATTTAAGTGTGTGTC	1200
Db	1141	AATAACTGTTCCACAGGGGCACTGCGAGGACCTGGTTAACGGATTTAAGTGTGTGTC	1200
QY	1201	CCCCCAGTGGACTGGGAAAACGTGCGAGTTAGATGCAAAATGAAATGAGGGCAAACT	1260
Db	1201	CCCCCAGTGGACTGGGAAAACGTGCGAGTTAGATGCAAAATGAAATGAGGGCAAACT	1260
QY	1261	TGTGTAAACGCCAAATCCTGTGAAAGTCTATTTGCCAGCTACTACTGGGACTGTTCCC	1320
Db	1261	TGTGTAAACGCCAAATCCTGTGAAAGTCTATTTGCCAGCTACTACTGGGACTGTTCCC	1320

QY	1321	GGCTGGATGGGTGAGAAATTTGATGATGCTTGGCAGTGTGCAAT	1380
Db	1321	GGCTGGATGGGTGAGAAATTTGATGATGCTTGGCAGTGTGCAAT	1380
QY	1381	GACGCTCTGTCGGGATTTGTTAATGTTATCGCTGTATCTGTCCACTGCTGTATGCA	1440
Db	1381	GACGCTCTGTCGGGATTTGTTAATGTTATCGCTGTATCTGTCCACTGCTGTATGCA	1440
QY	1441	GGCGATCATGTCGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGT	1500
Db	1441	GGCGATCATGTCGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGT	1500
QY	1501	CACCTGTCAGAAATCAACAGATTCAGTGTCTGTGCCCACTGTTTCTCGAAAC	1560
Db	1501	CACCTGTCAGAAATCAACAGATTCAGTGTCTGTGCCCACTGTTTCTCGAAAC	1560
QY	1561	CTCTGTGAGTGAATTCAGTATTTGTGAGCTTAATCCCTGCCAGAAACGGTGCAGTGC	1620
Db	1561	CTCTGTGAGTGAATTCAGTATTTGTGAGCTTAATCCCTGCCAGAAACGGTGCAGTGC	1620
QY	1621	TACAAACGCTGCCAGTACTATTTCTGCAAGTGCCTGAGGACTATGAGGCAAGAACTGC	1680
Db	1621	TACAAACGCTGCCAGTACTATTTCTGCAAGTGCCTGAGGACTATGAGGCAAGAACTGC	1680
QY	1681	TACACCTGAAAGACCACTGCCGACGACCCCTGTGAAAGTGAATGACAGCTGCACAGTG	1740
Db	1681	TACACCTGAAAGACCACTGCCGACGACCCCTGTGAAAGTGAATGACAGCTGCACAGTG	1740
QY	1741	GCATGGCTTCCAAACGACACCTGAAAGGCTGGGTATATTTCTTCCAAAGTCTGTGCT	1800
Db	1741	GCATGGCTTCCAAACGACACCTGAAAGGCTGGGTATATTTCTTCCAAAGTCTGTGCT	1800
QY	1801	CCTCAAGGAGTGCAGAGTCAAGTGCAGGAGGCAATTCACCTGTGACTGTGTAAACAGGC	1860
Db	1801	CCTCAAGGAGTGCAGAGTCAAGTGCAGGAGGCAATTCACCTGTGACTGTGTAAACAGGC	1860
QY	1861	TTCAACGGGAACTATCTGCCATGAAATATTAATGACTGTGAGAGCAACCCCTGTGAAAC	1920
Db	1861	TTCAACGGGAACTATCTGCCATGAAATATTAATGACTGTGAGAGCAACCCCTGTGAAAC	1920
QY	1921	GGTGGCACTTGCATCGATGGTCAACTCTCTCAAGTGCATCTGTAGTGAACGCTGGGAG	1980
Db	1921	GGTGGCACTTGCATCGATGGTCAACTCTCTCAAGTGCATCTGTAGTGAACGCTGGGAG	1980
QY	1981	GGGGCTTACTGTGAAACCAATATTAATGACTGTGAGAGCAACCCCTGCGCAATGGGGC	2040
Db	1981	GGGGCTTACTGTGAAACCAATATTAATGACTGTGAGAGCAACCCCTGCGCAATGGGGC	2040
QY	2041	ACGTGCGGACCTGGTCAATGACTTCTACTGTGACTGTAAAAATGGGTGGAAGGAAAG	2100
Db	2041	ACGTGCGGACCTGGTCAATGACTTCTACTGTGACTGTAAAAATGGGTGGAAGGAAAG	2100
QY	2101	ACCTGCACTCAAGTGCATGAGTGAAGGCAACGTCGCAACAAACGGTGGCACCTGC	2160
Db	2101	ACCTGCACTCAAGTGCATGAGTGAAGGCAACGTCGCAACAAACGGTGGCACCTGC	2160
QY	2161	TATGATGAGGGGATGCTTTTAAAGTGCATGTGTCTGGCGCTGGGAGGAAACAACTGT	2220
Db	2161	TATGATGAGGGGATGCTTTTAAAGTGCATGTGTCTGGCGCTGGGAGGAAACAACTGT	2220
QY	2221	AACATAGCCGAAACAGTGTCTGCTGCCCAACCCCTGCCATTAATGGGGGCAATGTGTG	2280
Db	2221	AACATAGCCGAAACAGTGTCTGCTGCCCAACCCCTGCCATTAATGGGGGCAATGTGTG	2280
QY	2281	GTCAAACGGGAGTCTTTTAAAGTGCATGAGGAGGCTGGGAGGGGCGCACTGTGTCT	2340
Db	2281	GTCAAACGGGAGTCTTTTAAAGTGCATGAGGAGGCTGGGAGGGGCGCACTGTGTCT	2340
QY	2341	CAGATACCAATGACTGACGCTTCTATCCCTGTTTACAAAGCGGCACTGTGTGTGGA	2400
Db	2341	CAGATACCAATGACTGACGCTTCTATCCCTGTTTACAAAGCGGCACTGTGTGTGGA	2400
QY	2401	GACAACTGGTACCGGTGCGAAATGTGCCCGGGTTTTTGTCTGGGCGCGGACTGCGAAATAAC	2460

181 CCGGAGACCGAAGTGCACCCGCGAGCGAGTGTGACACATACCTTCAAAGTGTCCCTCAAG 240
182 |||||
594 CCGGGAGACCGAAGTGCACCCGCGAGCGAGTGTGACACATACCTTCAAAGTGTCCCTCAAG 653
595 |||||
241 GAGTATCAGTCCGCGGTCAAGCGCGCGGGGCGCTTGCAGCTTCCGCTCAGGGTCCAGCGCT 300
242 |||||
654 GAGTATCAGTCCGCGGTCAAGCGCGCGGGGCGCTTGCAGCTTCCGCTCAGGGTCCAGCGCT 713
655 |||||
301 GTCATCGGGGGCAACACCTTCAACCTCAAGCGCAGCGCGGGCAACGACCGCAACCGGATC 360
302 |||||
714 GTCATCGGGGGCAACACCTTCAACCTCAAGCGCAGCGCGGGCAACGACCGCAACCGGATC 773
715 |||||
361 GTCTGTCTTTCAGTTCGCTCGCGAGGTCTCTATACGTTGCTTGTGAGGGGTGGGAT 420
362 |||||
774 GTCTGTCTTTCAGTTCGCTCGCGAGGTCTCTATACGTTGCTTGTGAGGGGTGGGAT 833
775 |||||
421 TCCAGTAATGACACCGTTCAACCTGACAGTATTAATGAAAGGCTTCTCACTCGGGCATG 480
422 |||||
834 TCCAGTAATGACACCGTTCAACCTGACAGTATTAATGAAAGGCTTCTCACTCGGGCATG 893
835 |||||
481 ATCAACCCCGAGCGGAGTGGCAGACGCTGAAGCAGAACACCGGCGTTCGCCACTTTGAG 540
482 |||||
894 ATCAACCCCGAGCGGAGTGGCAGACGCTGAAGCAGAACACCGGCGTTCGCCACTTTGAG 953
895 |||||
541 TATCAGATCCGCGTGACCTGTGATGACTACTACTATGCTTTGGCTGTYAATAAGTTCTGC 600
542 |||||
954 TATCAGATCCGCGTGACCTGTGATGACTACTACTATGCTTTGGCTGCAATAAGTTCTGC 1013
955 |||||
601 CGCCCCAGAGATGACTTCTTTGGACACTATGCTCTGTGACAGAAATGGCAACAAACTTGC 660
602 |||||
1014 CGCCCCAGAGATGACTTCTTTGGACACTATGCTCTGTGACAGAAATGGCAACAAACTTGC 1073
1015 |||||
661 ATGGAAGGCTGGATGGGCGCCGAATGTAACAGAGCTATTTTGGCAGCAAGGCTGGAGTCT 720
662 |||||
1074 ATGGAAGGCTGGATGGGCGCCGAATGTAACAGAGCTATTTTGGCAGCAAGGCTGGAGTCT 1133
1075 |||||
721 AAGCATGGGTCTTGCAAACTCCAGGTGACTGCGAGGTGCCAGTAYGGCTGCGCAAGGCTG 780
722 |||||
1134 AAGCATGGGTCTTGCAAACTCCAGGTGACTGCGAGGTGCCAGTATGGCTGGCAAGGCTG 1193
1135 |||||
781 TACTGTGATAAGTGCATCCACACCCGGGATGGCTCCACCGGCATCTGTAATGAGCCCTGG 840
782 |||||
1194 TACTGTGATAAGTGCATCCACACCCGGGATGGCTCCACCGGCATCTGTAATGAGCCCTGG 1253
1195 |||||
841 CAGTGCCTCTGTGAGACCAACTGGGGCGCGCAGCTCTGTGACAAAAGATCTCAATTAAGT 900
842 |||||
1254 CAGTGCCTCTGTGAGACCAACTGGGGCGCGCAGCTCTGTGACAAAAGATCTCAATTAAGT 1313
1255 |||||
901 GGGACTCATCAGCCGTGTCTCAACGGGGGAACCTTGTAGCAACACAGGCGCTTGACAAATAT 960
902 |||||
1314 GGGACTCATCAGCCGTGTCTCAACGGGGGAACCTTGTAGCAACACAGGCGCTTGACAAATAT 1373
1315 |||||
961 CAGTGTTCCTGCCCTGAGGGGTATTACAGGACCCAACTGTGAAATTTGTGAGCAGCGCTGC 1020
962 |||||
1374 CAGTGTTCCTGCCCTGAGGGGTATTACAGGACCCAACTGTGAAATTTGTGAGCAGCGCTGC 1433
1375 |||||
1021 CTCTCTGATCCCTGTCAACACAGAGGACGCTGTAAAGGAGACCTCCCTGGGCTTTGAGTGT 1080
1022 |||||
1434 CTCTCTGATCCCTGTCAACACAGAGGACGCTGTAAAGGAGACCTCCCTGGGCTTTGAGTGT 1493
1435 |||||
1081 GAGTGTTCCTGCCCTGAGCGGCGCCACATGCTCTACAACATTTGATGACTGTCTCTCCT 1140
1082 |||||
1494 GAGTGTTCCTGCCCTGAGCGGCGCCACATGCTCTACAACATTTGATGACTGTCTCTCCT 1553
1495 |||||
1141 AATAACTGTTCCCAACGGGGGCACTTGCAGGACCTGGTTAAACGGAATTTAAGTGTGTGTC 1200
1142 |||||
1554 AATAACTGTTCCCAACGGGGGCACTTGCAGGACCTGGTTAAACGGAATTTAAGTGTGTGTC 1613
1555 |||||
1201 CCCCACAGTGGACTGGGAAAAAGTGGCAGTTAGATGCAATTAAGTGTGAGGGCCAAACCT 1260
1202 |||||
1614 CCCCACAGTGGACTGGGAAAAAGTGGCAGTTAGATGCAATTAAGTGTGAGGGCCAAACCT 1673
1615 |||||

1261 TGTGTAAACGCCAAATCTGTGAAGATCTCATTTGCCAGCTACTACTGCGACTGTCTTTCCC 1320
1262 |||||
1674 TGTGTAAACGCCAAATCTGTGAAGATCTCATTTGCCAGCTACTACTGCGACTGTCTTTCCC 1733
1675 |||||
1321 GGTGTGATGGGTGAGAAATTTGTGACATAAATATTAATGACTGCTTGGCCAGTGTGAGAT 1380
1322 |||||
1734 GGTGTGATGGGTGAGAAATTTGTGACATAAATATTAATGACTGCTTGGCCAGTGTGAGAT 1793
1735 |||||
1381 GAGCCCTCTGTGCGGATTTGGTTAATGTTATCGCTGTATCTGTCCACTGTGCTATGCA 1440
1382 |||||
1794 GAGCCCTCTGTGCGGATTTGGTTAATGTTATCGCTGTATCTGTCCACTGTGCTATGCA 1853
1795 |||||
1441 GGCATCACTGTGAGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGGT 1500
1442 |||||
1854 GGCATCACTGTGAGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGGT 1913
1855 |||||
1501 CACTGTCAAGATGAATAATCAACAGATTCAGTGTCTGTGTCTGTCCACTGTGTTCTCGAATC 1560
1502 |||||
1914 CACTGTCAAGATGAATAATCAACAGATTCAGTGTCTGTGTCTGTCCACTGTGTTCTCGAATC 1973
1915 |||||
1561 CTCTGTCACTGGACATCGATTAATTTGTGAGCCTTAATCCCTGCCAGAAACGGTGGCCAGTGC 1620
1562 |||||
1974 CTCTGTCACTGGACATCGATTAATTTGTGAGCCTTAATCCCTGCCAGAAACGGTGGCCAGTGC 2033
1975 |||||
1621 TACAACCGTCCCAAGTACTATTTCTGCAAGTGTCCCGAGGACTATGAGGCAAGAACTGC 1680
1622 |||||
2034 TACAACCGTCCCAAGTACTATTTCTGCAAGTGTCCCGAGGACTATGAGGCAAGAACTGC 2093
2035 |||||
1681 TCACACTGAAAGACACTGCGCAGACGCCCTGTGAAAGTGAATGACAGCTGCACAGTG 1740
1682 |||||
2094 TCACACTGAAAGACACTGCGCAGACGCCCTGTGAAAGTGAATGACAGCTGCACAGTG 2153
2095 |||||
1741 GCCATGGCTTCCAAACGACACACCTGAAGGGTGGGTATATTTCTCCAAACGCTGTGGT 1800
1742 |||||
2154 GCCATGGCTTCCAAACGACACACCTGAAGGGTGGGTATATTTCTCCAAACGCTGTGGT 2213
2155 |||||
1801 CCTCAGGGGAAGTGCAGAGTCACTCGGGAGGCAAAATTCACCTGTGACTGTATCAAAGGC 1860
1802 |||||
2214 CCTCAGGGGAAGTGCAGAGTCACTCGGGAGGCAAAATTCACCTGTGACTGTATCAAAGGC 2273
2215 |||||
1861 TTCAACGGGAACATACCTGCCATGAAATATTAATGACTGTGAGAGCAACCCCTGTAGAAAC 1920
1862 |||||
2274 TTCAACGGGAACATACCTGCCATGAAATATTAATGACTGTGAGAGCAACCCCTGTAGAAAC 2333
2275 |||||
1921 GGTGGCACTTGCATCGATGGTGTCAACTCTCAAGTGCATCTGTAGTGCAGCTGGAG 1980
1922 |||||
2334 GGTGGCACTTGCATCGATGGTGTCAACTCTCAAGTGCATCTGTAGTGCAGCTGGAG 2393
2335 |||||
1981 GGGGCTTACTGTGAAACCAATATTAATGACTGCAGCCAGAAACCCCTGCCAATGGGGC 2040
1982 |||||
2394 GGGGCTTACTGTGAAACCAATATTAATGACTGCAGCCAGAAACCCCTGCCAATGGGGC 2453
2395 |||||
2041 ACCTGTGCGAACCTGTGCAATGACTTCTACTGTGACTGTGATAAATAATGGGTGGAAGGAAG 2100
2042 |||||
2454 ACCTGTGCGAACCTGTGCAATGACTTCTACTGTGACTGTGATAAATAATGGGTGGAAGGAAG 2513
2455 |||||
2101 ACCTGCCACTCAGTGTGACAGTGTGAGCCACAGTGCACCAACAGCTGGCACCTGC 2160
2102 |||||
2514 ACCTGCCACTCAGTGTGACAGTGTGAGCCACAGTGCACCAACAGCTGGCACCTGC 2573
2515 |||||
2161 TATGATGAGGGGATGCTTTTAAAGTGCATGTCTGTGGCGGCTGGGAAGGAACAACTGT 2220
2162 |||||
2574 TATGATGAGGGGATGCTTTTAAAGTGCATGTCTGTGGCGGCTGGGAAGGAACAACTGT 2633
2575 |||||
2221 AATAGCCCGAAACAGTAGCTGCTGCCCAAACCCCTGCCATTAATGGGGGCACTGTGTG 2280
2222 |||||
2634 AATAGCCCGAAACAGTAGCTGCTGCCCAAACCCCTGCCATTAATGGGGGCACTGTGTG 2693
2635 |||||
2281 GTCAACGGGCACTCTTTTACGTGCTGCAAGGAAGGCTGGGAGGGGCCCATCTGTGCT 2340
2282 |||||
2694 GTCAACGGGCACTCTTTTACGTGCTGCAAGGAAGGCTGGGAGGGGCCCATCTGTGCT 2753
2695 |||||
2341 CAGAAATCAATGACTGACGCCCTCATCTCTGTTTACAACAGCGGCACTGTGTGGATGGA 2400
2342 |||||

Db 2754 CAGAAATACCAATGACTGACGACCTCATCCCTGTTTACAAACGCGGCACCTGTGTGGATGGA 2813
Qy 2401 GACAACTGGTACCGGTGCGAATGTGCCCGGGTTTTCTGTGGGCCGAGCTGCGAATAAAC 2460
Db 2814 GACAACTGGTACCGGTGCGAATGTGCCCGGGTTTTCTGTGGGCCGAGCTGCGAATAAAC 2873
Qy 2461 ATCAATGAATGCCAGTCTTCACTTGTGCTTTGGAGCGGACCTGTGTGGATGATCAAT 2520
Db 2874 ATCAATGAATGCCAGTCTTCACTTGTGCTTTGGAGCGGACCTGTGTGGATGATCAAT 2933
Qy 2521 GGCTACCGGTGTGCTGCCCTCCAGGGGACAGTGGTCCCAAGTCCAGGAAGTTTCAGGG 2580
Db 2934 GGCTACCGGTGTGCTGCCCTCCAGGGGACAGTGGTCCCAAGTCCAGGAAGTTTCAGGG 2993
Qy 2581 AGACCTTGCATCACCATGGGGAGTGTGATCCAGATGGGGCCCAATGGGATGATGACTGT 2640
Db 2994 AGACCTTGCATCACCATGGGGAGTGTGATCCAGATGGGGCCCAATGGGATGATGACTGT 3053
Qy 2641 AATACCTGCCAGTCCCTGAATGACGGATCGCTGTCTCAAGGTCTGGTGTGGCCCTCGA 2700
Db 3054 AATACCTGCCAGTCCCTGAATGACGGATCGCTGTCTCAAGGTCTGGTGTGGCCCTCGA 3113
Qy 2701 CCTTGCCTGCTCCACAAAGGGCAGCGAGTGGCCAGCGGGCAGAGCTGCATCCCATC 2760
Db 3114 CCTTGCCTGCTCCACAAAGGGCAGCGAGTGGCCAGCGGGCAGAGCTGCATCCCATC 3173
Qy 2761 CTGGAGCAGCAGTCTTCTGCTCCACCCCTGCACCTGTGTGGCGGAGTGTCTGCTTCAGT 2820
Db 3174 CTGGAGCAGCAGTCTTCTGCTCCACCCCTGCACCTGTGTGGCGGAGTGTCTGCTTCAGT 3233
Qy 2821 CTCAGCGGTGAAGACAAAGTGACCTCTGACTCTCTATTTACCAGGATAACTGTGCGAAC 2880
Db 3234 CTCAGCGGTGAAGACAAAGTGACCTCTGACTCTCTATTTACCAGGATAACTGTGCGAAC 3293
Qy 2881 ATCAGATTACCTTTAACAGGAGATGATCTACCGAGTCTTACTACGGAGCATTTGC 2940
Db 3294 ATCAGATTACCTTTAACAGGAGATGATCTACCGAGTCTTACTACGGAGCATTTGC 3353
Qy 2941 AGTGAATTGAGGAATTTGAATATTTTGAAGAAATTTTCCGCTGAAATATTCATCTACATC 3000
Db 3354 AGTGAATTGAGGAATTTGAATATTTTGAAGAAATTTTCCGCTGAAATATTCATCTACATC 3413
Qy 3001 GCTTGGAGCTTTCCCTTTCAGGAAACAAATGAATAATCATGTGCGCCATTTCTGCTGAAGAT 3060
Db 3414 GCTTGGAGCTTTCCCTTTCAGGAAACAAATGAATAATCATGTGCGCCATTTCTGCTGAAGAT 3473
Qy 3061 ATACGGGATGATGGGAAACCGATCAAGGAATCACTGACAAAATATTCATCTGTTAGT 3120
Db 3474 ATACGGGATGATGGGAAACCGATCAAGGAATCACTGACAAAATATTCATCTGTTAGT 3533
Qy 3121 AAACGTGATGAAACAGCTCGCTGATGCTGCGGTTGCGAAGTAAAGATTACAGGCGG 3180
Db 3534 AAACGTGATGAAACAGCTCGCTGATGCTGCGGTTGCGAAGTAAAGATTACAGGCGG 3593
Qy 3181 CCTCTGAAGAACAGAAACAGATTTCTTGTTCCTTGTCTGAGCTCTGTCTTAACTGTGGCT 3240
Db 3594 CCTCTGAAGAACAGAAACAGATTTCTTGTTCCTTGTCTGAGCTCTGTCTTAACTGTGGCT 3653
Qy 3241 TGGATCTGTGCTGTGAGCGGCTTCTACTGTGCTGCGGAAAGCGGCGGAGCGCGG 3300
Db 3654 TGGATCTGTGCTGTGAGCGGCTTCTACTGTGCTGCGGAAAGCGGCGGAGCGCGG 3713
Qy 3301 AGCCACACACACTCAGCTCTGAGGACCAACACCAACCAACCAACCAACCAACCAACCAAC 3360
Db 3714 AGCCACACACACTCAGCTCTGAGGACCAACACCAACCAACCAACCAACCAACCAACCAAC 3773
Qy 3361 CAGATCAAAAACCCCAATTGAGAAACATGGGGCCCAACACCGTCCCTCATCAAGGATTACGAG 3420
Db 3774 CAGATCAAAAACCCCAATTGAGAAACATGGGGCCCAACACCGTCCCTCATCAAGGATTACGAG 3833
Qy 3421 AACAGAACTCCAAATGTCTAAATTAAGGACACAAATTTCTGAAGTAGAAGAGGACGAC 3480

Db 3834 AACAGAACTCCAAATGTCTAAATAAGGACACACAAATTTCTGAAGTAGAAGAGGACGAC 3893
Qy 3481 ATGGAACAAACACACAGAGAAAGCCCGGTTTGGCAAGCAGCGCGGTATAGCTCGTAGAC 3540
Db 3894 ATGGAACAAACACACAGAGAAAGCCCGGTTTGGCAAGCAGCGCGGTATAGCTCGTAGAC 3953
Qy 3541 AGAGAGAGAGAGGCCCCCAACCGGACCGCCGACCAAAACACCAAACTGGACAAACAAACAG 3600
Db 3954 AGAGAGAGAGAGGCCCCCAACCGGACCGCCGACCAAAACACCAAACTGGACAAACAAACAG 4013
Qy 3601 GACAAACAGAGACTTGGAAAGTCCAGAGCTTAAACCGAATGAGTACATCGTATAG 3657
Db 4014 GACAAACAGAGACTTGGAAAGTCCAGAGCTTAAACCGAATGAGTACATCGTATAG 4070

RESULT 3

US-09-566-047-1
; Sequence 1, Application US/09566047
; Patent No. 6703198
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; Hood, Leroy
; Krantz, Ian D.
; Spinner, Nancy B.
; TITLE OF INVENTION: Methods of Diagnosing Alagille Syndrome
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/566,047
; FILING DATE: 05-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/882,046
; FILING DATE: 25-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UW 4164
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (858) 535-9001
; TELEFAX: (858) 535-8949
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5590 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 414..4068
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-566-047-1

Query Match 99.9%; Score 3654.2; DB 3; Length 5590;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3653; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGCGTTCCCAACGAGACGCGCGGTCGCGGCGCCCTTAAGCTCTGCTCGCCCTG 60
Db 414 ATGCGTTCCCAACGAGACGCGCGGTCGCGGCGCCCTTAAGCTCTGCTCGCCCTG 473

Qy	61	CTCTGTGCGCTCGAGCCAGGTGTGTGGGCGCTCGGGTCAGTTTCAGATTGGAGATCCTG	120
Db	474	CTCTGTGCGCTCGAGCCAGGTGTGTGGGCGCTCGGGTCAGTTTCAGATTGGAGATCCTG	533
Qy	121	TCATGCGAGAACGTGAACCGGGAGCTGCGAGAACCGGAACTGTGCGCGCGCGCGCGGAAAC	180
Db	534	TCATGCGAGAACGTGAACCGGGAGCTGCGAGAACCGGAACTGTGCGCGCGCGCGCGGAAAC	593
Qy	181	CCGGGAGACCGAAGTGTGACCCCGGAGAGTGTGACATACATCTCAAAAGTGTGCTCAAG	240
Db	594	CCGGGAGACCGAAGTGTGACCCCGGAGAGTGTGACATACATCTCAAAAGTGTGCTCAAG	653
Qy	241	GAGTATCAGTCCCGCGTCAAGCGCGGGGGCGCTGCAAGCTTCAGGCTCAGGGTCCAGCGCT	300
Db	654	GAGTATCAGTCCCGCGTCAAGCGCGGGGGCGCTGCAAGCTTCAGGCTCAGGGTCCAGCGCT	713
Qy	301	GTATCGGGGGCAACCTTTCAACCTCAAGGCCAGCGCGGCAACCGCAACCGCATC	360
Db	714	GTATCGGGGGCAACCTTTCAACCTCAAGGCCAGCGCGGCAACCGCAACCGCATC	773
Qy	361	GTGCTGCTTTTCAAGTTCGGCTGCGCGAGGTCTATACGTGTGCTGTGGAGGGCGTGGAT	420
Db	774	GTGCTGCTTTTCAAGTTCGGCTGCGCGAGGTCTATACGTGTGCTGTGGAGGGCGTGGAT	833
Qy	421	TCCAGTAATGACACCGCTTCAACCTGACAGTATTATTGAAAAGCTTCTCACTCGGGCATG	480
Db	834	TCCAGTAATGACACCGCTTCAACCTGACAGTATTATTGAAAAGCTTCTCACTCGGGCATG	893
Qy	481	ATCAACCCCGAGCGGAGTGGCAGCGCTGAAGCAGAACCGGGCGTGTGCCCATCTTTGAG	540
Db	894	ATCAACCCCGAGCGGAGTGGCAGCGCTGAAGCAGAACCGGGCGTGTGCCCATCTTTGAG	953
Qy	541	TATCAGATCCGGCTGACTGTGATGACTACTATATGGCTTTGGCTGTAATAATGTTCTGC	600
Db	954	TATCAGATCCGGCTGACTGTGATGACTACTATATGGCTTTGGCTGTAATAATGTTCTGC	1013
Qy	601	CGCCCCAGAGATGACTTTTGGACACTATGCCCTGTGACAGAAATGGCAACAAACTTGC	660
Db	1014	CGCCCCAGAGATGACTTTTGGACACTATGCCCTGTGACAGAAATGGCAACAAACTTGC	1073
Qy	661	ATGGAAGGCTGGATGGGCGCGAATGTAAACAGAGCTATTGTCGACAAAGGCTCGAGTCCT	720
Db	1074	ATGGAAGGCTGGATGGGCGCGAATGTAAACAGAGCTATTGTCGACAAAGGCTCGAGTCCT	1133
Qy	721	AAGCATGGCTTGTGCAAACTCCAGGTGACTGAGGTGACAGGTGACAGGTGACAGGTGAC	780
Db	1134	AAGCATGGCTTGTGCAAACTCCAGGTGACTGAGGTGACAGGTGACAGGTGACAGGTGAC	1193
Qy	781	TACTGTGATAAGTGCATCCACACCCGGAGTGGTCCACGGCATCTGTAATGAGCCCTGG	840
Db	1194	TACTGTGATAAGTGCATCCACACCCGGAGTGGTCCACGGCATCTGTAATGAGCCCTGG	1253
Qy	841	CAGTGGCTCTGTGAGACCAACTGGGGCGGCGAGCTCTGTGACAAAGATCTCAATTACTGT	900
Db	1254	CAGTGGCTCTGTGAGACCAACTGGGGCGGCGAGCTCTGTGACAAAGATCTCAATTACTGT	1313
Qy	901	GGGACTCATAGCGGTGTCTCAACGGGGGAACTTTGTAGCAACACAGGCGCTGACAAATAT	960
Db	1314	GGGACTCATAGCGGTGTCTCAACGGGGGAACTTTGTAGCAACACAGGCGCTGACAAATAT	1373
Qy	961	CAGTGTTCCTGCGCTGAGGGTATTTCAGGACCCCAACTGTGAAATTTGCTGAGCAGCGCTGC	1020
Db	1374	CAGTGTTCCTGCGCTGAGGGTATTTCAGGACCCCAACTGTGAAATTTGCTGAGCAGCGCTGC	1433
Qy	1021	CTCTCTGATCCCTGTCAACACAGAGGCAAGCTGTGAAGGAGACCTCCCTGGGCTTTGAGTGT	1080
Db	1434	CTCTCTGATCCCTGTCAACACAGAGGCAAGCTGTGAAGGAGACCTCCCTGGGCTTTGAGTGT	1493
Qy	1081	GAGTGTTCCTGCGCTGAGGGTATTTCAGGACCCCAACTGTGAAATTTGCTGAGCAGCGCTGC	1140
Db	1494	GAGTGTTCCTGCGCTGAGGGTATTTCAGGACCCCAACTGTGAAATTTGCTGAGCAGCGCTGC	1553
Qy	1141	AATAACTGTTCACCGGGGCACTCTGCCAGGACCTGGTTAACGGATTAAAGTGTGTGTG	1200

Db	1554	AATAACTGTTCACCGGGGCACTCTGCCAGGACCTGGTTAAGGATTTAAGTGTGTGTGC	1613
Qy	1201	CCCCCAGTGCAGCTGGGGAACCGTCCAGTTAGATGCAAAATGAATGTGAGGCAAACT	1260
Db	1614	CCCCCAGTGCAGCTGGGGAACCGTCCAGTTAGATGCAAAATGAATGTGAGGCAAACT	1673
Qy	1261	TGTGTAAACCGCAAAATCCCTGTAAAGATCTCATTTGCCAGCTACTACTGCGACCTGTCTTCCC	1320
Db	1674	TGTGTAAACCGCAAAATCCCTGTAAAGATCTCATTTGCCAGCTACTACTGCGACCTGTCTTCCC	1733
Qy	1321	GGCTGATGGGTGAGAAATTTGTGACATAAATAATTAATGACTGCTTGGCCAGGTGACAAAT	1380
Db	1734	GGCTGATGGGTGAGAAATTTGTGACATAAATAATTAATGACTGCTTGGCCAGGTGACAAAT	1793
Qy	1381	GACGCTCTGCTCGGATTTGGTTAATGTTATCGCTGTATCTGTCCACTGCTGCTATGCA	1440
Db	1794	GACGCTCTGCTCGGATTTGGTTAATGTTATCGCTGTATCTGTCCACTGCTGCTATGCA	1853
Qy	1441	GGCGATCACTGTGAGAGACATCCGATGAATGTGCCAGCAACCCCTGTGTTGAAATGGGGT	1500
Db	1854	GGCGATCACTGTGAGAGACATCCGATGAATGTGCCAGCAACCCCTGTGTTGAAATGGGGT	1913
Qy	1501	CACGTGCAGAAATGAAATCAACAGATTCAGATGCTGTGTGCCACTGGTTTCTCTGAAAC	1560
Db	1914	CACGTGCAGAAATGAAATCAACAGATTCAGATGCTGTGTGCCACTGGTTTCTCTGAAAC	1973
Qy	1561	CTCTGTGAGTGCAGATTCGATTTTGTGAGCCTTAATCCCTGCCAGAACCGTGCCAGTGC	1620
Db	1974	CTCTGTGAGTGCAGATTCGATTTTGTGAGCCTTAATCCCTGCCAGAACCGTGCCAGTGC	2033
Qy	1621	TACAAACCGTGCAGTGAATTTCTGCAAGTGTGCCAGGACTATGAGGGCAAGAACTGC	1680
Db	2034	TACAAACCGTGCAGTGAATTTCTGCAAGTGTGCCAGGACTATGAGGGCAAGAACTGC	2093
Qy	1681	TACACCTGAAAGACCACTGCCAGCAGCCCTGTGAAAGTGAATGACAGCTGCACAGTG	1740
Db	2094	TACACCTGAAAGACCACTGCCAGCAGCCCTGTGAAAGTGAATGACAGCTGCACAGTG	2153
Qy	1741	GCCATGGCTTCCAAACGACACCTGGAAGGGTGGGTATATTTCTCCCAACGCTGTGGT	1800
Db	2154	GCCATGGCTTCCAAACGACACCTGGAAGGGTGGGTATATTTCTCCCAACGCTGTGGT	2213
Qy	1801	CCTCAGGGAAAGTGAAGAGTCAAGTGGGAGGCAAAATTCACCTGTGACTGTAAACAAAGC	1860
Db	2214	CCTCAGGGAAAGTGAAGAGTCAAGTGGGAGGCAAAATTCACCTGTGACTGTAAACAAAGC	2273
Qy	1861	TTCAACGGGAAACATATGCTGCCATGAAATATTAATGACTGTGAGAGCAACCCCTGTGAAAC	1920
Db	2274	TTCAACGGGAAACATATGCTGCCATGAAATATTAATGACTGTGAGAGCAACCCCTGTGAAAC	2333
Qy	1921	GGTGGCACTTGCATCGATGGTGTCAACTCTCAAGTGCATCTGTAGTGACGGCTGGAG	1980
Db	2334	GGTGGCACTTGCATCGATGGTGTCAACTCTCAAGTGCATCTGTAGTGACGGCTGGAG	2393
Qy	1981	GGGCGCTTACTGTGAAACCAATATTAATGACTGTGAGCAGCAACCCCTGCCCAATGGGGC	2040
Db	2394	GGGCGCTTACTGTGAAACCAATATTAATGACTGTGAGCAGCAACCCCTGCCCAATGGGGC	2453
Qy	2041	ACGTGTGCGACCTGGTCAATGACTTCTACTGTGACTGTGAAATGGGTGAAAGGAAAG	2100
Db	2454	ACGTGTGCGACCTGGTCAATGACTTCTACTGTGACTGTGAAATGGGTGAAAGGAAAG	2513
Qy	2101	ACCTGCCACTCAGCTGACAGTGTGATGAGCCAGCTGCAACACCGTGGCACCTGC	2160
Db	2514	ACCTGCCACTCAGCTGACAGTGTGATGAGCCAGCTGCAACACCGTGGCACCTGC	2573
Qy	2161	TATGATCAGGGGATGCTTTTAAAGTGCATGTGCTCGCGCTGGGAGGAAACCACTGT	2220
Db	2574	TATGATCAGGGGATGCTTTTAAAGTGCATGTGCTCGCGCTGGGAGGAAACCACTGT	2633
Qy	2221	AACATAGCCGAAACAGTGTGCTGCCCAACCCCTGCCCAATTAATGGGGGCAATGTGTG	2280

Db 714 GT CATCGGGGCAACACCTTTCAACCTCAAGCCAGCGCGCAACGACCGCAATC 773
Qy 361 GTCTGCTTTTCAAGTTTCGGCTGGCGAGGTCTTATACGTTGCTTGTGGAGGCTGGAT 420
Db 774 GTCTGCTTTTCAAGTTTCGGCTGGCGAGGTCTTATACGTTGCTTGTGGAGGCTGGAT 833
Qy 421 TCCAGTAATGACACCGTTCAACCTGACAGTATTATTGAAAAGGCTTCTCACTCGGCGATG 480
Db 834 TCCAGTAATGACACCGTTCAACCTGACAGTATTATTGAAAAGGCTTCTCACTCGGCGATG 893
Qy 481 ATCAACCCAGCGCGAGTGGCAGAGCTGAAGCAGAACACGGCGTGTGCCACTTTGAG 540
Db 894 ATCAACCCAGCGCGAGTGGCAGAGCTGAAGCAGAACACGGCGTGTGCCACTTTGAG 953
Qy 541 TATCAGATCCGCGTGACTGTGATGACTACTACTATGCGCTTTGGCTGATTAAGTTCTGC 600
Db 954 TATCAGATCCGCGTGACTGTGATGACTACTACTATGCGCTTTGGCTGCAATAAGTTCTGC 1013
Qy 601 CGCCCCAGAGATGACTTCTTTGGACACTATGCTGTGACAGAAATGGCAACAAAATTCG 660
Db 1014 CGCCCCAGAGATGACTTCTTTGGACACTATGCTGTGACAGAAATGGCAACAAAATTCG 1073
Qy 661 ATGGAAGGCTGGATGGGCCCGGAATGTAACAGAGCTATTGTCCGACAAAGCTGCAATCCT 720
Db 1074 ATGGAAGGCTGGATGGGCCCGGAATGTAACAGAGCTATTGTCCGACAAAGCTGCAATCCT 1133
Qy 721 AAGCATGGGCTTGGCAACTCCAGGTGACTGCAAGTGGCAGTAYGGCTGGCAGGCGCTG 780
Db 1134 AAGCATGGGCTTGGCAAACTCCAGGTGACTGCAAGTGGCAGTAYGGCTGGCAGGCGCTG 1193
Qy 781 TACTGTGATAAGTGATCTCCACACCCCGGATGCGTCCACGGCATCTGTAATGAGCCCTGG 840
Db 1194 TACTGTGATAAGTGATCTCCACACCCCGGATGCGTCCACGGCATCTGTAATGAGCCCTGG 1253
Qy 841 CAGTGCCCTCTGTGAGACCAACTGGGGCGGCAGCTCTGTGACAAAAGATCTCAATTAAGTGT 900
Db 1254 CAGTGCCCTCTGTGAGACCAACTGGGGCGGCAGCTCTGTGACAAAAGATCTCAATTAAGTGT 1313
Qy 901 GGGACTCATCAGCCGCTGTCTCAAGCGGGGAACCTGTGAGCAACACAGGCCCTGCAAAATAT 960
Db 1314 GGGACTCATCAGCCGCTGTCTCAAGCGGGGAACCTGTGAGCAACACAGGCCCTGCAAAATAT 1373
Qy 961 CAGTGTTCCTGCTCAGGGGTATTTCAAGGACCCAACTGTGAAATGTGTGAGCAGCCCTGC 1020
Db 1374 CAGTGTTCCTGCTCAGGGGTATTTCAAGGACCCAACTGTGAAATGTGTGAGCAGCCCTGC 1433
Qy 1021 CTCTCTGATCCCTGTCAACACAGAGGCAAGCTGTGAAGGAGACCTCCCTGGGCTTTGAGTGT 1080
Db 1434 CTCTCTGATCCCTGTCAACACAGAGGCAAGCTGTGAAGGAGACCTCCCTGGGCTTTGAGTGT 1493
Qy 1081 GAGTGTTCCTCAGGCTGGACCGGCCCCACATGCTCTACAAACATTGATGACTGTCTCCT 1140
Db 1494 GAGTGTTCCTCAGGCTGGACCGGCCCCACATGCTCTACAAACATTGATGACTGTCTCCT 1553
Qy 1141 AATAACTGTTCCACGGGGCACTGCGCAGGACCTGGTTAACGGATTTAAAGTGTGTGTC 1200
Db 1554 AATAACTGTTCCACGGGGCACTGCGCAGGACCTGGTTAACGGATTTAAAGTGTGTGTC 1613
Qy 1201 CCCCACAGTGGACTGGGAAAACGTGCGAGTTAGATGCAATGAATGTGAGGCGCAAACT 1260
Db 1614 CCCCACAGTGGACTGGGAAAACGTGCGAGTTAGATGCAATGAATGTGAGGCGCAAACT 1673
Qy 1261 TGTGTAAACGCCAAATCTGTGAAGATCTAATTCGCGAGCTACTACTCGGACTGTCTCCC 1320
Db 1674 TGTGTAAACGCCAAATCTGTGAAGATCTAATTCGCGAGCTACTACTCGGACTGTCTCCC 1733
Qy 1321 GGCTGGATGGGTGAGAAATTTGACATAAATATTAAATGACTGCTTGGCCAGTGTGAGAAAT 1380
Db 1734 GGCTGGATGGGTGAGAAATTTGACATAAATATTAAATGACTGCTTGGCCAGTGTGAGAAAT 1793
Qy 1381 GACGCTCTCTGTCGGGATTTGGTTAATGGTTATCGCTGTATCTGTCCACCTGGCTATGCA 1440
Db 1794 GACGCTCTCTGTCGGGATTTGGTTAATGGTTATCGCTGTATCTGTCCACCTGGCTATGCA 1853

Qy 1441 GCGATCACTGTGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGT 1500
Db 1854 GCGATCACTGTGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGT 1913
Qy 1501 CACTGTGAGAAATGAATCAACAGATTCAGTGTCTGTGTCCACACTGGTTTCTCTGAAAC 1560
Db 1914 CACTGTGAGAAATGAATCAACAGATTCAGTGTCTGTGTCCACACTGGTTTCTCTGAAAC 1973
Qy 1561 CTCTGTGAGTGGACATCGATTTATTGTGAGCTTAATCCCTGCCAGAACGGTGCCTAGTGC 1620
Db 1974 CTCTGTGAGTGGACATCGATTTATTGTGAGCTTAATCCCTGCCAGAACGGTGCCTAGTGC 2033
Qy 1621 TACAAACCGTGCAGTCACTATTCTCAAGTGGCCCGAGAGCTATGAGGGCAAGAACTGC 1680
Db 2034 TACAAACCGTGCAGTCACTATTCTTCAAGTGGCCCGAGAGCTATGAGGGCAAGAACTGC 2093
Qy 1681 TCACACTGAAAGACCACTGCCCGACGACCCCTGTGAAAGTATTGACAGCTGCACAGTG 1740
Db 2094 TCACACTGAAAGACCACTGCCCGACGACCCCTGTGAAAGTATTGACAGCTGCACAGTG 2153
Qy 1741 GCCATGGCTTCCAAACGACACACTGGAAGGTGCGGTATATTTCTCCACAGCTCTGTGT 1800
Db 2154 GCCATGGCTTCCAAACGACACACTGGAAGGTGCGGTATATTTCTCCACAGCTCTGTGT 2213
Qy 1801 CCTCAGCGAAAGTGCAGAGTCACTCGGGAGGCAAAATTCACCTGTGACTGTGTAAACAAAGGC 1860
Db 2214 CCTCAGCGAAAGTGCAGAGTCACTCGGGAGGCAAAATTCACCTGTGACTGTGTAAACAAAGGC 2273
Qy 1861 TTCACGGGAAACATPACTGCCAATATAATTAATGACTGTGAGAGCAACCCCTGTGTAGAAC 1920
Db 2274 TTCACGGGAAACATPACTGCCAATATAATTAATGACTGTGAGAGCAACCCCTGTGTAGAAC 2333
Qy 1921 GGTGGCACTTGCATCGATGGTGTCAACTCTCAAGTGGATCTGTAGTGAAGCTGGAG 1980
Db 2334 GGTGGCACTTGCATCGATGGTGTCAACTCTCTCAAGTGGATCTGTAGTGAAGCTGGAG 2393
Qy 1981 GGGGCTTACTGTGAAACCAATATAATTAATGACTGACGACGCAACCCCTGCCCACAATGGGGGC 2040
Db 2394 GGGGCTTACTGTGAAACCAATATAATTAATGACTGACGACGCAACCCCTGCCCACAATGGGGGC 2453
Qy 2041 ACCTGTGCGACCTGTGCAATGACTCTTACTGTGACTGTGTAATAATGGGTGGAAGAAAG 2100
Db 2454 ACCTGTGCGACCTGTGCAATGACTCTTACTGTGACTGTGTAATAATGGGTGGAAGAAAG 2513
Qy 2101 ACCTGCCACTCACTGACAGTGTGTGAGGCGCACGTCGCAACACGCTGACCTGC 2160
Db 2514 ACCTGCCACTCACTGACAGTGTGTGAGGCGCACGTCGCAACACGCTGACCTGC 2573
Qy 2161 TATGATGAGGGGATGCTTTTAAAGTGCATGTGTCTTGGCGGCTGGGAAGGAAACAACCTGT 2220
Db 2574 TATGATGAGGGGATGCTTTTAAAGTGCATGTGTCTTGGCGGCTGGGAAGGAAACAACCTGT 2633
Qy 2221 AACATAGCCGAAACAGTGTGCTGCTCCCAACCCCTGCCATTAATGGGGGCAATGTGTG 2280
Db 2634 AACATAGCCGAAACAGTGTGCTGCTCCCAACCCCTGCCATTAATGGGGGCAATGTGTG 2693
Qy 2281 GTCAACGGCGAGTCTTTACGTGGTGTCAAGGAGGCTGGGAGGGGCCCATCTGTGCT 2340
Db 2694 GTCAACGGCGAGTCTTTTACGTGGTGTCAAGGAGGCTGGGAGGGGCCCATCTGTGCT 2753
Qy 2341 CAGAAATACCAATGACTGCAGCCCTCATCCCTGTTTACAAACAGCGGACCTGTGTGGATGGA 2400
Db 2754 CAGAAATACCAATGACTGCAGCCCTCATCCCTGTTTACAAACAGCGGACCTGTGTGGATGGA 2813
Qy 2401 GACAACTGGTACCGGTGCGAAATGTGCCCGGGTTTTCTGGGCCCGGACTGCGAGAAATAAC 2460
Db 2814 GACAACTGGTACCGGTGCGAAATGTGCCCGGGTTTTCTGGGCCCGGACTGCGAGAAATAAC 2873
Qy 2461 ATCAATGAATGCCAGTCTTCACTTGTGCTTTGGAGGCGACCTGTGTGGATGAGATCAAT 2520
Db 2874 ATCAATGAATGCCAGTCTTCACTTGTGCTTTGGAGGCGACCTGTGTGGATGAGATCAAT 2933

|||||
Db 949 TATCAGATCCGGTGACCTGTGATGACTACTACTAATAAGCTTTGGCTGCAATAAGTTCTGC 1008
Qy 601 CGCCCCAGAGATGACTTTCTTTTGGACACTATGCCCTGTGACACAGAAATGGCAACAAACTTGC 660
Db 1009 CGCCCCAGAGATGACTTTCTTTGGACACTATGCCCTGTGACACAGAAATGGCAACAAACTTGC 1068
Qy 661 ATGGAAGGCTGGATGGGCCCGCAATGTAAACAGAGCTATTTGCCGACAAAGGCTCGAGTCCCT 720
Db 1069 ATGGAAGGCTGGATGGGCCCGCAATGTAAACAGAGCTATTTGCCGACAAAGGCTCGAGTCCCT 1128
Qy 721 AAGCATGGGCTTCTGCAAACTCCCAAGTGACTGACAGTGGCCAGTAYGGCTGGCAAGGCTTG 780
Db 1129 AAGCATGGGCTTCTGCAAACTCCCAAGTGACTGACAGTGGCCAGTAYGGCTGGCAAGGCTTG 1188
Qy 781 TACTGTGATAGTGCAATCCCAACCCCGGATGGCTCACCGGCATCTGTAAATGAGCCCTGG 840
Db 1189 TACTGTGATAGTGCAATCCCAACCCCGGATGGCTCACCGGCATCTGTAAATGAGCCCTGG 1248
Qy 841 CAGTGCCTCTGTGAGACCAACTGGGGCGGCAGCTCTGTGACAAAGATCTCAATTAAGTGT 900
Db 1249 CAGTGCCTCTGTGAGACCAACTGGGGCGGCAGCTCTGTGACAAAGATCTCAATTAAGTGT 1308
Qy 901 GGGACTCATCAGCCGTGTCTCAACGGGGGAACCTTGTAGCAACACAGGCCCTGTGACAAATAT 960
Db 1309 GGGACTCATCAGCCGTGTCTCAACGGGGGAACCTTGTAGCAACACAGGCCCTGTGACAAATAT 1368
Qy 961 CAGTGTCTGCCCTGAGGGGTATTTAGGACCCCAACTGTGAAATTTGTGAGCAGCCCTGC 1020
Db 1369 CAGTGTCTGCCCTGAGGGGTATTTAGGACCCCAACTGTGAAATTTGTGAGCAGCCCTGC 1428
Qy 1021 CTCTCTGATCCCTGTCAACAGAGGAGCTGTGTAAGAGACCTCCCTGGGCTTTGAGTGT 1080
Db 1429 CTCTCTGATCCCTGTCAACAGAGGAGCTGTGTAAGAGACCTCCCTGGGCTTTGAGTGT 1488
Qy 1081 GAGTGTCTCCAGAGCTGGACGGGCCACACATGCTCTACAAACATTTGATGACTGTCTTCCT 1140
Db 1489 GAGTGTCTCCAGAGCTGGACGGGCCACACATGCTCTACAAACATTTGATGACTGTCTTCCT 1548
Qy 1141 AATAACTGTCTCCAGGGGGCACTGCGAGACCTGGTTAAACCGATTTAAAGTGTGTGC 1200
Db 1549 AATAACTGTCTCCAGGGGGCACTGCGAGACCTGGTTAAACCGATTTAAAGTGTGTGC 1608
Qy 1201 CCCCCACAGTGGACTGGGAAACGTGCCAGTTAGATGCAATGATGTGAGGCCAAACCT 1260
Db 1609 CCCCCACAGTGGACTGGGAAACGTGCCAGTTAGATGCAATGATGTGAGGCCAAACCT 1668
Qy 1261 TGTGTAAACGCCAAATCTCTGAAGAATCTCATTTGCCAGCTACTACTGCGACTGTCTTCCC 1320
Db 1669 TGTGTAAACGCCAAATCTCTGAAGAATCTCATTTGCCAGCTACTACTGCGACTGTCTTCCC 1728
Qy 1321 GGCTGGATGGGTGAGAAATTTGACATAAATAATTAATGATGCTGCTTGGCCAGTGTGAGAAAT 1380
Db 1729 GGCTGGATGGGTGAGAAATTTGACATAAATAATTAATGATGCTGCTTGGCCAGTGTGAGAAAT 1788
Qy 1381 GACGCTCTCTGCGGATTTGGTTAATGGTTATCGCTGTATCTGTCCACCTGGCTATGCA 1440
Db 1789 GACGCTCTCTGCGGATTTGGTTAATGGTTATCGCTGTATCTGTCCACCTGGCTATGCA 1848
Qy 1441 GCGGATCACTGTGAGAGACATCGATGAATGTGCAGCAACCCCTGTTTCAATGGGGGT 1500
Db 1849 GCGGATCACTGTGAGAGACATCGATGAATGTGCAGCAACCCCTGTTTCAATGGGGGT 1908
Qy 1501 CACTGTGAGAAATGAATCAACAGATTCACAGTGTCTGTGTCCCACTGGTTTCTCTGGAAC 1560
Db 1909 CACTGTGAGAAATGAATCAACAGATTCACAGTGTCTGTGTCCCACTGGTTTCTCTGGAAC 1968
Qy 1561 CTCTGTGAGTGGACATCGATTTATGTGAGCTTAATCCCTGCCAGACGGTGGCCAGTGC 1620
Db 1969 CTCTGTGAGTGGACATCGATTTATGTGAGCTTAATCCCTGCCAGACGGTGGCCAGTGC 2028
Qy 1621 TACAACCGTGCAGTGAATAATTTCTGCAAGTGGCCCGAGGACTATGAGGGCAAGAACTGC 1680
|||||

2029 TACAACCGTGCAGTGACTATTTCTGCAAGTGGCCCGAGGACTATGAGGGCAAGAACTGC 2088
1681 TCACACCTGAAAGACCACTGCCGACGACCCCTGTGAAGTGAATTCACAGCTGCACAGTG 1740
2089 TCACACCTGAAAGACCACTGCCGACGACCCCTGTGAAGTGAATTCACAGCTGCACAGTG 2148
1741 GCCATGGCTTCCAAACGACACACCTGAAGGGGTGCGGTATATTTCTCTCCCAACGCTCTGGT 1800
2149 GCCATGGCTTCCAAACGACACACCTGAAGGGGTGCGGTATATTTCTCTCCCAACGCTCTGGT 2208
1801 CTTCAACGGGAAGTGCAGAGTCAAGAGTCAAGTGGAGGCAAAATTCACCTGTGATGTATCAAAAGGC 1860
2209 CTTCAACGGGAAGTGCAGAGTCAAGAGTCAAGTGGAGGCAAAATTCACCTGTGATGTATCAAAAGGC 2268
1861 TTTCAACGGGAACATACCTGCCATGAAATATTAATGACTGTGAGAGCAACCCCTGTGAGAAAC 1920
2269 TTTCAACGGGAACATACCTGCCATGAAATATTAATGACTGTGAGAGCAACCCCTGTGAGAAAC 2328
1921 GGTGGCACTTGCATCGATGGTGTCAACTCTCAAGTGCATCTGTAGTGAACCGCTGGAG 1980
2329 GGTGGCACTTGCATCGATGGTGTCAACTCTCTACAGTGCATCTGTAGTGAACCGCTGGAG 2388
1981 GGGGCTTACTGTGAAACCAATATTAATGACTGCAGCCAGAAACCCCTGCGCACAATGGGGGC 2040
2389 GGGGCTTACTGTGAAACCAATATTAATGACTGCAGCCAGAAACCCCTGCGCACAATGGGGGC 2448
2041 ACCTGTGCGGACCTGTGCAATGACTTCTACTGTGACTGTAAATAATGGTGGAAAGAAAG 2100
2449 ACCTGTGCGGACCTGTGCAATGACTTCTACTGTGACTGTAAATAATGGTGGAAAGAAAG 2508
2101 ACCTGCCACTCAGTGCAGTCAAGTGTGATGAGGCCACGTCGCAACCAACCGTGGCACCTGC 2160
2509 ACCTGCCACTCAGTGCAGTCAAGTGTGATGAGGCCACGTCGCAACCAACCGTGGCACCTGC 2568
2161 TATGATGAGGGGATGCTTTTAAAGTGCATGTCTCTGGCGGCTGGGAGGAAACCAACTGT 2220
2569 TATGATGAGGGGATGCTTTTAAAGTGCATGTCTCTGGCGGCTGGGAGGAAACCAACTGT 2628
2221 AACATAGCCGAAACAGTAGTCCCTGCCCAACCCCTGCCATTAATGGGGCACATGTGTG 2280
2629 AACATAGCCGAAACAGTAGTCCCTGCCCAACCCCTGCCATTAATGGGGCACATGTGTG 2688
2281 GTCAACGGCGAGTCTTTTACGTGCTCTGCAAGGAAGGCTGGGAGGGGGCCCATCTGTGCT 2340
2689 GTCAACGGCGAGTCTTTTACGTGCTCTGCAAGGAAGGCTGGGAGGGGGCCCATCTGTGCT 2748
2341 CAGAATAACCAATGATGACGCCCTCATCTCTGTTTACAAACAGCGGCACCTGTGTGGATGGA 2400
2749 CAGAATAACCAATGATGACGCCCTCATCTCTGTTTACAAACAGCGGCACCTGTGTGGATGGA 2808
2401 GACAACTGGTACCGGTGCGAATGTGCCCGGGTTTTGCTGGGCCGACTGCAGAAATAAC 2460
2809 GACAACTGGTACCGGTGCGAATGTGCCCGGGTTTTGCTGGGCCGACTGCAGAAATAAC 2868
2461 ATCAATGAATGCCAGTCTTCACTTGTGCTCTTTGGAGCGACCTGTGTGATGATCAAT 2520
2869 ATCAATGAATGCCAGTCTTCACTTGTGCTCTTTGGAGCGACCTGTGTGATGATCAAT 2928
2521 GGCTACCGGTGTGTCTCCCTCCAGGGCAAGTGGTGGCCAGTGGCCAGGAGTTCAGGG 2580
2929 GGCTACCGGTGTGTCTCCCTCCAGGGCAAGTGGTGGCCAGTGGCCAGGAGTTCAGGG 2988
2581 AGACCTTGCAATCACCATTGGGAGTGTATACAGATGGGGCCCAATGGGATGATGACTGT 2640
2989 AGACCTTGCAATCACCATTGGGAGTGTATACAGATGGGGCCCAATGGGATGATGACTGT 3048
2641 AATACCTGCCAGTGCCTGAATGGAGGATCGCTGTCTCAAAGGTCTGTGTGGCCCTCGA 2700
3049 AATACCTGCCAGTGCCTGAATGGAGGATCGCTGTCTCAAAGGTCTGTGTGGCCCTCGA 3108
2701 CTTTGCCTGTCTCAAAAAGGGCACAGAGTGGCCCGAGGGGAGAGCTGCATCCCCATC 2760
3109 CTTTGCCTGTCTCAAAAAGGGCACAGAGTGGCCCGAGGGAGAGCTGCATCCCCATC 3168
|||||

QY	841	CAGTGCTCTGTGAGCCAACTGGGGCGGCAGCTCTGTGACAAAGATCTCAATTA	900
Db	1253	CAGTGCCCTCTGTGAGACCAACTGGGGCGGCAGCTCTGTGACAAAGATCTCAATTA	1312
QY	901	GGGACTCATCAGCGCTGTCTCAACGGGGGAACCTTTGTAGCAACACAGGCCCTTGACAAATAT	960
Db	1313	GGGACTCATCAGCGCTGTCTCAACGGGGGAACCTTTGTAGCAACACAGGCCCTTGACAAATAT	1372
QY	961	CAGTGTTCCTGCCCTGAGGGGTATTTAGGACCCAACTGTGTGAATTTGCTGAGCAAGCCTGC	1020
Db	1373	CAGTGTTCCTGCCCTGAGGGGTATTTAGGACCCAACTGTGTGAATTTGCTGAGCAAGCCTGC	1432
QY	1021	CTCTCTGATCCCTGTGACACAGAGGAGCAGCTGTAAAGGAGACCTCCCTGGGCTTTGAGTGT	1080
Db	1433	CTCTCTGATCCCTGTGACACAGAGGAGCAGCTGTAAAGGAGACCTCCCTGGGCTTTGAGTGT	1492
QY	1081	GAGTGTTCCTGCCCTGAGCGGCCCCACATGTCTCTACAAACATTTGATGACTGTCTCCT	1140
Db	1493	GAGTGTTCCTGCCCTGAGCGGCCCCACATGTCTCTACAAACATTTGATGACTGTCTCCT	1552
QY	1141	AATAACTGTTTCCACGGGGGACCTGCCAGGACTGGTTTAAACCGAATTTAAGTGTGTGC	1200
Db	1553	AATAACTGTTTCCACGGGGGACCTGCCAGGACTGGTTTAAACCGAATTTAAGTGTGTGC	1612
QY	1201	CCCCCAGTGGACTGGGAAAAAGTGCAGTTAGATGCAATGATGAGGCGCAACCT	1260
Db	1613	CCCCCAGTGGACTGGGAAAAAGTGCAGTTAGATGCAATGATGAGGCGCAACCT	1672
QY	1261	TGTGTAAACCGCAATCCTGTGAAGATCTCATTTGCCAGTACTACTGCGACTGTCTTCCC	1320
Db	1673	TGTGTAAACCGCAATCCTGTGAAGATCTCATTTGCCAGTACTACTGCGACTGTCTTCCC	1732
QY	1321	GGCTGAGTGGGTGAGAAATTTGACATAAAATATTAATGACTGCCTTGGCCAGTGTGAGAAAT	1380
Db	1733	GGCTGAGTGGGTGAGAAATTTGACATAAAATATTAATGACTGCCTTGGCCAGTGTGAGAAAT	1792
QY	1381	GAGCCCTCTGTGCGGATTTGGTTAAATGGTTATCGCTGTATCTGTCCACCTGGCTATGCA	1440
Db	1793	GAGCCCTCTGTGCGGATTTGGTTAAATGGTTATCGCTGTATCTGTCCACCTGGCTATGCA	1852
QY	1441	GGCGATCACTGTGAGAGAGACATCGATGAATGTGCCAGCAACCCCTCTTTGAATGGGGGT	1500
Db	1853	GGCGATCACTGTGAGAGAGACATCGATGAATGTGCCAGCAACCCCTCTTTGAATGGGGGT	1912
QY	1501	CACGTGCAGATGAATCAACAGATTCAGTGTCTGTGTGCCACTGTGTTCTCTGGAAC	1560
Db	1913	CACGTGCAGATGAATCAACAGATTCAGTGTCTGTGTGCCACTGTGTTCTCTGGAAC	1972
QY	1561	CTCTGTGAGTGGACATCGATTAATTTGTAGGCTTAATCCCTGCCAGACGGTGGCCAGTGC	1620
Db	1973	CTCTGTGAGTGGACATCGATTAATTTGTAGGCTTAATCCCTGCCAGACGGTGGCCAGTGC	2032
QY	1621	TACAACTGTGCAGTACTATTTCTGCAAGTGGCCCGAGGACTATGAGGGCAAGAACTGC	1680
Db	2033	TACAACTGTGCAGTACTATTTCTGCAAGTGGCCCGAGGACTATGAGGGCAAGAACTGC	2092
QY	1681	TCACACCTGAAAGACCACTGCCGACGACCCCTGTGTGAAGTGTGACAGCTGCACAGTG	1740
Db	2093	TCACACCTGAAAGACCACTGCCGACGACCCCTGTGTGAAGTGTGACAGCTGCACAGTG	2152
QY	1741	GCCATGGCTTCCACGACACACTGAAAGGGGTGCGGTATATTTCTCCAGAGTCTGTGGT	1800
Db	2153	GCCATGGCTTCCACGACACACTGAAAGGGGTGCGGTATATTTCTCCAGAGTCTGTGGT	2212
QY	1801	CCTCACGGGAAGTGAAGAGTCACTGCGGGAGGCAAAATTCACCTGTGACTGTAAACAAAGGC	1860
Db	2213	CCTCACGGGAAGTGAAGAGTCACTGCGGGAGGCAAAATTCACCTGTGACTGTAAACAAAGGC	2272
QY	1861	TTACAGGGAAACATCTGCGCATGAAATAATTAATGACTGTGTGAGAGCAACCCCTGTGAGAAC	1920
Db	2273	TTACAGGGAAACATCTGCGCATGAAATAATTAATGACTGTGTGAGAGCAACCCCTGTGAGAAC	2332

QY	1921	GGTGGCACTTGATCGATGGTGTCAACTCTCTACAAAGTGCATCTGTAGTGACCGCTGGGAG	1980
Db	2333	GGTGGCACTTGATCGATGGTGTCAACTCTCTACAAAGTGCATCTGTAGTGACCGCTGGGAG	2392
QY	1981	GGGGCCCTACTGTGAAACCAATATTAATGACTGAGGAGCAGAACCCCTGCCACAAATGGGGC	2040
Db	2393	GGGGCCCTACTGTGAAACCAATATTAATGACTGAGGAGCAGAACCCCTGCCACAAATGGGGC	2452
QY	2041	AGTGTGCGGAGCCTGGTCAATGACTTCTACTGTGACTGTAAAAATGGGTGGAAGGAAAG	2100
Db	2453	AGTGTGCGGAGCCTGGTCAATGACTTCTACTGTGACTGTAAAAATGGGTGGAAGGAAAG	2512
QY	2101	ACCTGCCACTCAGCTGACAGTGTGATGAGGCGCACCTGTGCAACCAACCGTGGCACCTGC	2160
Db	2513	ACCTGCCACTCAGCTGACAGTGTGATGAGGCGCACCTGTGCAACCAACCGTGGCACCTGC	2572
QY	2161	TATGATGAGGGGATGCTTTTAAAGTGCATGTGTCTGCGGGCTGTGGGAGGAAACAACTGT	2220
Db	2573	TATGATGAGGGGATGCTTTTAAAGTGCATGTGTCTGCGGGCTGTGGGAGGAAACAACTGT	2632
QY	2221	AACATAGCCGGAACAGTAGCTGCCCAACCCCTGCCATAATGGGGGACATGTGTG	2280
Db	2633	AACATAGCCGGAACAGTAGCTGCCCAACCCCTGCCATAATGGGGGACATGTGTG	2692
QY	2281	GTCAACGGCGAGTCTTTTACGTGCGTCTGCAAGGAAAGGCTGGGAGGGGCCCATCTGTGCT	2340
Db	2693	GTCAACGGCGAGTCTTTTACGTGCGTCTGCAAGGAAAGGCTGGGAGGGGCCCATCTGTGCT	2752
QY	2341	CAGAAATACCAATGACTGACAGCCCTCATCCCTGTTTACACAGCGGCACTGTGTGTGATGGA	2400
Db	2753	CAGAAATACCAATGACTGACAGCCCTCATCCCTGTTTACACAGCGGCACTGTGTGTGATGGA	2812
QY	2401	GACAACTGGTACCGGTGCGAAATGTGCCCGGGTTTCTGGGCGCCGACATGACAGAAATAAC	2460
Db	2813	GACAACTGGTACCGGTGCGAAATGTGCCCGGGTTTCTGGGCGCCGACATGACAGAAATAAC	2872
QY	2461	ATCAATGAATGCGAGTCTTTCACCTTGTGCGCTTTGGAGCGACCTGTGTGGATGAGATCAAT	2520
Db	2873	ATCAATGAATGCGAGTCTTTCACCTTGTGCGCTTTGGAGCGACCTGTGTGGATGAGATCAAT	2932
QY	2521	GGCTACCGGTGTGTGCGCTTCCAGGGCAAGTGTGTGCGCAAGTGCACGAGAGTTTCAGGG	2580
Db	2933	GGCTACCGGTGTGTGCGCTTCCAGGGCAAGTGTGTGCGCAAGTGCACGAGAGTTTCAGGG	2992
QY	2581	AGACCTTGATCACCATGGGGAGTGTATACAGATGGGGCCAAATGGGATGATGACTGT	2640
Db	2993	AGACCTTGATCACCATGGGGAGTGTATACAGATGGGGCCAAATGGGATGATGACTGT	3052
QY	2641	AATACCTGCGAGTGCCTGAAATGACCGGATCGCCTGTCTCAAAGGTCTGTGTGGCCCTCGA	2700
Db	3053	AATACCTGCGAGTGCCTGAAATGACCGGATCGCCTGTCTCAAAGGTCTGTGTGGCCCTCGA	3112
QY	2701	CCTTGGCTGTCCACAAAGGGCAAGAGTGTGCCAGCGGGCAGAGCTGCATCCCCATC	2760
Db	3113	CCTTGGCTGTCCACAAAGGGCAAGAGTGTGCCAGCGGGCAGAGCTGCATCCCCATC	3172
QY	2761	CTGAGCAGACAGTCTTCTGCTCCACCTGCACTGTGTGGGCGAGTGTCTTCCAGT	2820
Db	3173	CTGAGCAGACAGTCTTCTGCTCCACCTGCACTGTGTGGGCGAGTGTCTTCCAGT	3232
QY	2821	CTCAGCGCGTGAAGCAAAAGTGCACCTCTGACTCTCTATTTACCAGGATAAATCTGTGCGAAC	2880
Db	3233	CTCAGCGCGTGAAGCAAAAGTGCACCTCTGACTCTCTATTTACCAGGATAAATCTGTGCGAAC	3292
QY	2881	ATCACATTTTACCTTTTAAAGGAGATGATGTACAGGCTCTTACTACGGAGCATTGTGC	2940
Db	3293	ATCACATTTTACCTTTTAAAGGAGATGATGTACAGGCTCTTACTACGGAGCATTGTGC	3352
QY	2941	AGTGAATTTGAGGAATTTTGAATATTTTGAAGAAATGTTTCCCGCTGAATATTTCAATCTACATC	3000
Db	3353	AGTGAATTTGAGGAATTTTGAATATTTTGAAGAAATGTTTCCCGCTGAATATTTCAATCTACATC	3412
QY	3001	GCTTTCGAGCGCTTCCCTTTCAGCGGAAACAAATGAGAAATACATGTGGCCATTTTCTGTGGAAGAT	3060


```
Db 3413 GCTTGCAGGCTTCCCTTTCAGCGAACAAATGAATAATCATGTGGCCATTTCTGCTGAAGAT 3472
Qy 3061 ATACGGGATGATGGGAACCCGATCAAGCAAAATCACTGACAAAATAATTCGATCTTGTAGT 3120
Db 3473 ATACGGGATGATGGGAACCCGATCAAGCAAAATCACTGACAAAATAATTCGATCTTGTAGT 3532
Qy 3121 AAACGTGATGAAACAGCTCGCTGATTTGCTGCGGTTGCGAGTAAGAGTTCAAGAGCGG 3180
Db 3533 AAACGTGATGAAACAGCTCGCTGATTTGCTGCGGTTGCGAGTAAGAGTTCAAGAGCGG 3592
Qy 3181 CCTCTGAAGAACAGAAACAGATTTCTTGTTCCTTCTGCTGAGCTCTGTCTTAACCTGGCT 3240
Db 3593 CCTCTGAAGAACAGAAACAGATTTCTTGTTCCTTCTGCTGAGCTCTGTCTTAACCTGGCT 3652
Qy 3241 TGGATCTGTTGCTTGGTGAGCGGCTTCTACTGTTGCTGCGGAAGCGGCGGAAGCGGCGG 3300
Db 3653 TGGATCTGTTGCTTGGTGAGCGGCTTCTACTGTTGCTGCGGAAGCGGCGGAAGCGGCGG 3712
Qy 3301 AGCCACACACACTCAGCTCTGAGGACAAACCAACCAACAGCTGCGGAGCAGCTGAAC 3360
Db 3713 AGCCACACACACTCAGCTCTGAGGACAAACCAACCAACAGCTGCGGAGCAGCTGAAC 3772
Qy 3361 CAGATCAAAAACCCCATTTGAGAACATGGGGCCCAACAGCTGCCCATCAAGGATTACGAG 3420
Db 3773 CAGATCAAAAACCCCATTTGAGAACATGGGGCCCAACAGCTGCCCATCAAGGATTATGAG 3832
Qy 3421 AACAGAACTCCAAATGTTCTAAATTAAGGACACACAATTTCTGAAGTAGAAGAGGACGAC 3480
Db 3833 AACAGAACTCCAAATGTTCTAAATTAAGGACACACAATTTCTGAAGTAGAAGAGGACGAC 3892
Qy 3481 ATGGACAAACACAGCAGAGAACGCCGTTTGGCAAGCAGCGCGGTATACGCTGGTAGAC 3540
Db 3893 ATGGACAAACACAGCAGAGAACGCCGTTTGGCAAGCAGCGCGGTATACGCTGGTAGAC 3592
Qy 3541 AGAAGAGAGAGGCCCCCAAGCGGACGCCGACAAACCAACCAACCTGGACAAACAAACAG 3600
Db 3953 AGAAGAGAGAGGCCCCCAAGCGGACGCCGACAAACCAACCAACCTGGACAAACAAACAG 4012
Qy 3601 GACACAGAGACTTGGAAAGTGGCCAGAGCTTAAACCGAATGAGTACATCGTATAG 3657
Db 4013 GACACAGAGACTTGGAAAGTGGCCAGAGCTTAAACCGAATGAGTACATCGTATAG 4069
```

RESULT 7

```
US-09-917-254-34
; Sequence 34, Application US/09917254
; Patent No. 6703204
; GENERAL INFORMATION:
; APPLICANT: Mutter, George
; APPLICANT: Baak, Jan
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer
; FILE REFERENCE: B0801/7224(JRV)
; CURRENT APPLICATION NUMBER: US/09/917,254
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/222,093
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 4855
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-09-917-254-34
```

Query Match 99.8%; Score 3649.4; DB 3; Length 4855;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3650; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
Qy 1 ATCGTTTCCCGACGACACGCGCGGTCGGGGCCCTTAAGCCCTCCTGCTCGCCCTG 60
Db 371 ATCGTTTCCCGACGACACGCGCGGTCGGGGCCCTTAAGCCCTCCTGCTCGCCCTG 430
```

```
Qy 61 CTCTGTGCTCGCAGAGCAAGGTGTGTGGGGCTTCGGGTCAAGTTTCAGATTGGAGATCCTG 120
Db 431 CTCTGTGCTCGCAGAGCAAGGTGTGTGGGGCTTCGGGTCAAGTTTCAGATTGGAGATCCTG 490
Qy 121 TCCATGACAGAACTGAAACGGGGAGCTGCAGAAACGGGAACTGCTGCGGCGGCGCCGGAAC 180
Db 491 TCCATGACAGAACTGAAACGGGGAGCTGCAGAAACGGGAACTGCTGCGGCGGCGCCGGAAC 550
Qy 181 CCGGAGAGACCGCAAGTGCACCCCGGACGAGTGTGACACATATCTTCAAAGTTGTGCTCAAG 240
Db 551 CCGGAGAGACCGCAAGTGCACCCCGGACGAGTGTGACACATATCTTCAAAGTTGTGCTCAAG 610
Qy 241 GAGTATCAGTCCCGCTCAGCGCCGGGGGCGCTTCAGAGCTTCGAGCTCAGGGTTCACGCT 300
Db 611 GAGTATCAGTCCCGCTCAGCGCCGGGGGCGCTTCAGAGCTTCGAGCTCAGGGTTCACGCT 670
Qy 301 GTCATCGGGGGCAACACCTTCAACCTCAGGCCAGCGCGGCAACGACCGCAACCGCATC 360
Db 671 GTCATCGGGGGCAACACCTTCAACCTCAGGCCAGCGCGGCAACGACCGCAACCGCATC 730
Qy 361 GTGCTGCTTTTCAAGTTTTCGCTGCGCGAGGTCCCTATACGTTGCTTGTGAGGGCTGGAT 420
Db 731 GTGCTGCTTTTCAAGTTTTCGCTGCGCGAGGTCCCTATACGTTGCTTGTGAGGGCTGGAT 790
Qy 421 TCCAGTAATGACACCGTTCAACCTGACAGTATTTATGAAAGGCTTCTCACTCGGGCATG 480
Db 791 TCCAGTAATGACACCGTTCAACCTGACAGTATTTATGAAAGGCTTCTCACTCGGGCATG 850
Qy 481 ATCAACCCCAAGCGGAGTGGCAGCGCTGACGAGCAACGCGGCGTTCGCCATTTTGG 540
Db 851 ATCAACCCCAAGCGGAGTGGCAGCGCTGACGAGCAACGCGGCGTTCGCCATTTTGG 910
Qy 541 TATCAGATCCGCTGACCTGTGATGACTACTACTATGGCTTTTGGCTGTAATTAAGTTCTGC 600
Db 911 TATCAGATCCGCTGACCTGTGATGACTACTACTATGGCTTTTGGCTGTAATTAAGTTCTGC 970
Qy 601 CGCCCAGAGATGACTTTCTTTTGGACACTATGCTGTGACAGAAATGGCAACAAATTCG 660
Db 971 CGCCCAGAGATGACTTTCTTTTGGACACTATGCTGTGACAGAAATGGCAACAAATTCG 1030
Qy 661 ATGGAAGGCTGGATGGGGCCCGGAACTGTAACAGAGCTATTTGCCGACAAAGCTCAGTCCT 720
Db 1031 ATGGAAGGCTGGATGGGGCCCGGAACTGTAACAGAGCTATTTGCCGACAAAGCTCAGTCCT 1090
Qy 721 AAGCATGGGTCTTGCAAAACCTCCAGAGTGCAGGTGCAGTAYGGCTGGCAAGGCTG 780
Db 1091 AAGCATGGGTCTTGCAAAACCTCCAGAGTGCAGGTGCAGTAYGGCTGGCAAGGCTG 1150
Qy 781 TACTGTGATAAGTGCATCCACACCCCGGATGCTCCACGGCATCTGTAAATGAGCCCTGG 840
Db 1151 TACTGTGATAAGTGCATCCACACCCCGGATGCTCCACGGCATCTGTAAATGAGCCCTGG 1210
Qy 841 CAGTGTCTGTGAGACCAACTGGGGCGGCGAGCTCTGTGACAAAGATCTCAATTAAGTGT 900
Db 1211 CAGTGTCTGTGAGACCAACTGGGGCGGCGAGCTCTGTGACAAAGATCTCAATTAAGTGT 1270
Qy 901 GGGACTCATCAGCGCTGTCTCAACGGGGAACTTTGTAGCAACACAGCGCTTCGACAAATAT 960
Db 1271 GGGACTCATCAGCGCTGTCTCAACGGGGAACTTTGTAGCAACACAGCGCTTCGACAAATAT 1330
Qy 961 CAGTGTCTGTGCGCTCAGGGGTATTGAGCAACCACTGTGAAATGCTGTGAGCAGCGCTGC 1020
Db 1331 CAGTGTCTGTGCGCTCAGGGGTATTGAGCAACCACTGTGAAATGCTGTGAGCAGCGCTGC 1390
Qy 1021 CTCTCTGATCCCTGTGACACAGAGCGAGCTGTAAAGAGACCTCCCTGGGCTTTGAGTGT 1080
Db 1391 CTCTCTGATCCCTGTGACACAGAGCGAGCTGTAAAGAGACCTCCCTGGGCTTTGAGTGT 1450
Qy 1081 GAGTGTTCCTCCAGGCTGGACCGGCGCCCACTGCTTACAAAACATTTGATGATGCTTCTCCT 1140
Db 1451 GAGTGTTCCTCCAGGCTGGACCGGCGCCCACTGCTTACAAAACATTTGATGATGCTTCTCCT 1510
Qy 1141 AATAACTGTTTCCCAAGGGGGGACCTGTCAGGACCTGGTTAAACGGAATTTAAGTGTGTGC 1200
```


Db	1511	 AATAACTGTTCCCAACGGGGCACCCTGCAGGACCTGGTTAAACGGATTTAAGTGTGTGTGC	1570
Qy	1201	CCCCCAGTGGACTGGGAAAAAGTGCAGTTAGATGCAATGAAATGTGAGGCCAAACCT	1260
Db	1571	CCCCCAGTGGACTGGGAAAAAGTGCAGTTAGATGCAATGAAATGTGAGGCCAAACCT	1630
Qy	1261	TGTGTAAACGCCAAATCCTGTAAAGAACTCAATTGCCAGCTACTACTGCAGCTGTCTTCCC	1320
Db	1631	TGTGTAAACGCCAAATCCTGTAAAGAACTCAATTGCCAGCTACTACTGCAGCTGTCTTCCC	1690
Qy	1321	GGCTGGATGGGTGAGAAATTTGTGACATATAATTAATGACTGCCCTTGGCCAGTGTCAGAAT	1380
Db	1691	GGCTGGATGGGTGAGAAATTTGTGACATATAATTAATGACTGCCCTTGGCCAGTGTCAGAAT	1750
Qy	1381	GACGCCCTCTGTGGGATTTGGTTAAATGGTTATTCGCTGTATCTGTCCACTGCGCTATGCA	1440
Db	1751	GACGCCCTCTGTGGGATTTGGTTAAATGGTTATTCGCTGTATCTGTCCACTGCGCTATGCA	1810
Qy	1441	GGCGATCACTGTGAGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGGT	1500
Db	1811	GGCGATCACGTGAGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGGT	1870
Qy	1501	CACGTGCAGAAATGAATCAACAGATTCAGTGTCTGTGTCCCACTGGTTCCTCTGGAAC	1560
Db	1871	CACGTGCAGAAATGAATCAACAGATTCAGTGTCTGTGTCCCACTGGTTCCTCTGGAAC	1930
Qy	1561	CTCTGTGAGTGGACATCGATTATTTGTGAGCCTAATCCCTGCCAGAACGGTCCCACTGC	1620
Db	1931	CTCTGTGAGTGGACATCGATTATTTGTGAGCCTAATCCCTGCCAGAACGGTCCCACTGC	1990
Qy	1621	TACAACCGTGCAGTGACTATTTCTGCAAGTGCCCGAGGACTATGAGGCAAGAACTGC	1680
Db	1991	TACAACCGTGCAGTGACTATTTCTGCAAGTGCCCGAGGACTATGAGGCAAGAACTGC	2050
Qy	1681	TCACACCTGAAAGCACTGCGCGCACGCCCTGTCAAGTGAATGACAGCTGCACAGTG	1740
Db	2051	TCACACCTGAAAGCACTGCGCGCACGCCCTGTCAAGTGAATGACAGCTGCACAGTG	2110
Qy	1741	GCCATGGCTTCAACGACACACCTGAAGGGGTGGGTATATTTCTTCCACAGCTGTGGT	1800
Db	2111	GCCATGGCTTCAACGACACACCTGAAGGGGTGGGTATATTTCTTCCACAGCTGTGGT	2170
Qy	1801	CCTCACGGGAAGTCAAGAGTCAGTCGGGAGGCAAAATTCACCTGTGTACTGTAAACAAGGC	1860
Db	2171	CCTCACGGGAAGTCAAGAGTCAGTCGGGAGGCAAAATTCACCTGTGTACTGTAAACAAGGC	2230
Qy	1861	TTCAACGGGAACATATGCCATGAAAAATTAATGACTGTGAGAGCAACCCCTGTGAAAC	1920
Db	2231	TTCAACGGGAACATATGCCATGAAAAATTAATGACTGTGAGAGCAACCCCTGTGAAAC	2290
Qy	1921	GGTGGCACTTCGATCGATGGTGTCACTCCTACAAGTGCATCTGTGTAGTACGGCTGGAG	1980
Db	2291	GGTGGCACTTCGATCGATGGTGTCACTCCTACAAGTGCATCTGTGTAGTACGGCTGGAG	2350
Qy	1981	GGGGCTACTGTGAAAAACAAATTAATGACTGCAGCCAGAACCCCTGCCCAATGGGGGC	2040
Db	2351	GGGGCTACTGTGAAAAACAAATTAATGACTGCAGCCAGAACCCCTGCCCAATGGGGGC	2410
Qy	2041	ACGTGTCCGACCTGGTCAATGACTTCTACTGTGACTGTAAAAATGGTGGAAAGGAAAG	2100
Db	2411	ACGTGTCCGACCTGGTCAATGACTTCTACTGTGACTGTAAAAATGGTGGAAAGGAAAG	2470
Qy	2101	ACCTGCCACTCAGGTGACAGTGTGTGTATGAGGCCAGTGCACACGGTGGCACCTGC	2160
Db	2471	ACCTGCCACTCAGGTGACAGTGTGTGTATGAGGCCAGTGCACACGGTGGCACCTGC	2530
Qy	2161	TATGATGAGGGGATGCTTTTAAAGTGCAATGTCTCTGGCGCTGGGAAGGAAACACTGT	2220
Db	2531	TATGATGAGGGGATGCTTTTAAAGTGCAATGTCTCTGGCGCTGGGAAGGAAACACTGT	2590
Qy	2221	AACATAGCCGAAAAACAGTAGCTGCTGCTGCCAACCCCTGCCATAATGGGGGCACATGTGTG	2280

2591	AA	CATGCCCGAAACAGTAGCTGCTGCTGCCCAACCCCTTGCCATAAATGGGGGCA	CATGTGTG	2655
2281	GT	CAACGGCGAGTCCTTTAC	TGCGCTGTGCAAGGAAGGCTGGAGGGGCCCATCTGTGCT	2340
2651	GT	CAACGGCGAGTCCTTTAC	TGCGCTGTGCAAGGAAGGCTGGAGGGGCCCATCTGTGCT	2710
2341	CAGA	ATACCAATGACTGTGAGCCCTCAT	TCCCTGTTTAAACAGCGGCACCTGTGTGGATGGA	2400
2711	CAGA	ATACCAATGACTGTGAGCCCTCAT	TCCCTGTTTAAACAGCGGCACCTGTGTGGATGGA	2770
2401	GACA	CTGTGTACCGGTGCGAATGTG	CCCCGGGTTTTGCTGGGCCCGGACTGCGAATAAAC	2460
2771	GACA	CTGTGTGTACCGGTGCGAATGTG	CCCCGGGTTTTGCTGGGCCCGGACTGCGAATAAAC	2830
2461	ATCA	ATGAATGCCAGTCTTCA	CCCTGTGTGCTTTGGAGCGACCTGTGTGGATGAGATCAAT	2520
2831	ATCA	ATGAATGCCAGTCTTCA	CCCTGTGTGCTTTGGAGCGACCTGTGTGGATGAGATCAAT	2890
2521	GGCT	ACCGGTGTGTCTGCCCTC	CAGGGCAAGTGTGTGCCAAGTGCAGGGAAGTTTCAGGG	2580
2891	GGCT	ACCGGTGTGTCTGCCCTC	CAGGGCAAGTGTGTGCCAAGTGCAGGGAAGTTTCAGGG	2950
2581	AGAC	CTTGCATCACCATGGGAGTGT	ATACACAGATGGGGCCAAATGGGATGATGACTGT	2640
2951	AGAC	CTTGCATCACCATGGGAGTGT	ATACACAGATGGGGCCAAATGGGATGATGACTGT	3010
2641	AATAC	CTGCCAGTGCCTGAAATGGA	CGCGATCGCCTCAAAGGTCTGGTGTGGCCCTCGA	2700
3011	AATAC	CTGCCAGTGCCTGAAATGGA	CGCGATCGCCTCAAAGGTCTGGTGTGGCCCTCGA	3070
2701	CTTGT	CTGTCTCCAAAGGGCA	CAGCGATGCCCCAGCGGCAGAGCTGCATCCCCATC	2760
3071	CTTGT	CTGTCTCCAAAGGGCA	CAGCGATGCCCCAGCGGCAGAGCTGCATCCCCATC	3130
2761	CTGA	CGACCAAGTCTTCCTGCT	CCACCCCTGCACCTGTGTGGCGAGTGTCCGTCTTCAGT	2820
3131	CTGA	CGACCAAGTCTTCCTGCT	CCACCCCTGCACCTGTGTGGCGAGTGTCCGTCTTCAGT	3190
2821	CTCC	AGCGGTGAAGACAAAGTGAC	CTCTGACTCTTATACACAGGATAACTGTGCGAAC	2880
3191	CTCC	AGCGGTGAAGACAAAGTGAC	CTCTGACTCTTATACACAGGATAACTGTGCGAAC	3250
2881	ATCA	CATTTACTTTTAA	CAAGGAGATGTCTACCGGTCTTACTACGGAGCATTTGCT	2940
3251	ATCA	CATTTACTTTTAA	CAAGGAGATGTCTACCGGTCTTACTACGGAGCATTTGCT	3310
2941	AGTGA	ATTTGAGGAATTTTGAATAT	TTTTCAGAAATGTTTTCCGCTGAAATTTCAAATCTACATC	3000
3311	AGTGA	ATTTGAGGAATTTTGAATAT	TTTTCAGAAATGTTTTCCGCTGAAATTTCAAATCTACATC	3370
3001	GCTT	GCGAGCCTTCCCTT	CAGCGAAACAATGAATACTATGTGGCATTCTGTCTGAAGAT	3060
3371	GCTT	GCGAGCCTTCCCTT	CAGCGAAACAATGAATACTATGTGGCATTCTGTCTGAAGAT	3430
3061	ATAC	CGGATGATGGAAACCC	CGATCAAGGAAATCACTGACAAAATAATCGATCTTTGTTAGT	3120
3431	ATAC	CGGATGATGGAAACCC	CGATCAAGGAAATCACTGACAAAATAATCGATCTTTGTTAGT	3490
3121	AAA	CGTGTATGAAACAGCT	CGCTGATTTGCTGCCGTTGCAGAAATGAGTTTCAGAGCGG	3180
3491	AAA	CGTGTATGAAACAGCT	CGCTGATTTGCTGCCGTTGCAGAAATGAGTTTCAGAGCGG	3550
3181	CCTC	TGAAGAACAGAA	CAGATTTTCTTTGTTCCCTTGTCTGAGCTCTGTCTTAACTGTGGCT	3240
3551	CCTC	TGAAGAACAGAA	CAGATTTTCTTTGTTCCCTTGTCTGAGCTCTGTCTTAACTGTGGCT	3610
3241	TGGAT	CTGTTTGGTGA	CGCCCTTCTACTGGTGCTCTGCGGAAGCGCGAAGCCGGGC	3300
3611	TGGAT	CTGTTTGGTGA	CGCCCTTCTACTGGTGCTCTGCGGAAGCGCGAAGCCGGGC	3670
3301	AGCC	ACACACTCAG	CCCTCTGAGCAACACACCAACCAACACGTCGGGAGCAGCTGAAC	3360
3671	AGCC	ACACACTCAG	CCCTCTGAGCAACACACCAACCAACACGTCGGGAGCAGCTGAAC	3730

QY 3361 CAGATCAAAAACCCCATTTGAGAAACATGGGGCCAAACACGGTCCCATCAAGGATTACGAG 3420
DB 3731 CAGATCAAAAACCCCATTTGAGAAACATGGGGCCAAACACGGTCCCATCAAGGATTACGAG 3790
QY 3421 AACAGAACTCCAAATGCTCTAAATTAAGGACACACAATTTCTGAAGTAGAGAGGACGAC 3480
DB 3791 AACAGAACTCCAAATGCTCTAAATTAAGGACACACAATTTCTGAAGTAGAGAGGACGAC 3850
QY 3481 ATGACAAAACACCAAGAGAGAGCCCGGTTTGGCAGAGAGCCGGGTATACGCTGTTAGAC 3540
DB 3851 ATGACAAAACACCAAGAGAGAGCCCGGTTTGGCAGAGAGCCGGGTATACGCTGTTAGAC 3910
QY 3541 AG 3600
DB 3911 AG 3970
QY 3601 GACAAACAGAGACTTGGAAAGTGGCCAGAGCTTAAACCGAATGGAGTACATCGTATAG 3657
DB 3971 GACAAACAGAGACTTGGAAAGTGGCCAGAGCTTAAACCGAATGGAGTACATCGTATAG 4027

RESULT 8

US-09-214-278-6

; Sequence 6, Application US/09214278

; Patent No. 6291210

; GENERAL INFORMATION:

; APPLICANT: Sakano, Seiji

; APPLICANT: Itoh, Akira

; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE

; FILE REFERENCE: KP-8576

; CURRENT APPLICATION NUMBER: US/09/214, 278

; CURRENT FILING DATE: 1999-01-26

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 4208

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (409)..(4062)

; NAME/KEY: sig_peptide

; LOCATION: (409)..(501)

; NAME/KEY: mat_peptide

; LOCATION: (503)..(4062)

US-09-214-278-6

Query Match 99.8%; Score 3647.8; DB 3; Length 4208;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 3649; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGGTTTCCCACGACRCCGGCGGTTCGGGGCCCGCCCTAGCCCTCTGCTGGCCCTG 60
DB 409 ATGGTTTCCCACGACRCCGGCGGTTCGGGGCCCGCCCTAGCCCTCTGCTGGCCCTG 468
QY 61 CTCTGTCCCTCGCAGCAAGGTGTGTGGGCGCTTCGGGTTCAGATTGAGATCCCTG 120
DB 469 CTCTGTCCCTCGCAGCAAGGTGTGTGGGCGCTTCGGGTTCAGATTGAGATCCCTG 528
QY 121 TCCATGCAAGCTGAAACGGGGAGCTGCAGAAACGGGAACTGTCTGCGCGCGCGCCGGAAC 180
DB 529 TCCATGCAAGCTGAAACGGGGAGCTGCAGAAACGGGAACTGTCTGCGCGCGCGCCGGAAC 588
QY 181 CCGGGAGACCGAAGTGCACCCGGGAGGTGTGACATCTTCAAGTGTGCTCAAG 240
DB 589 CCGGGAGACCGAAGTGCACCCGGGAGGTGTGACATCTTCAAGTGTGCTCAAG 648
QY 241 GAGTATCAGTCCCGCGTCAAGCGCGGGGGCGCTTCGAGCTTCAGGGTCCAGCCCT 300
DB 649 GAGTATCAGTCCCGCGTCAAGCGCGGGGGCGCTTCGAGCTTCAGGGTCCAGCCCT 708
QY 301 GTCATCGGGGGCAACACCTTCAACCTCAAGGCCAGCGCGGCAACGACCGCAACCGCATC 360

DB 709 GTCATCGGGGGCAACACCTTCAACCTCAAGGCCAGCGCGCAACGACCGCAACCGCATC 768
QY 361 GTGCTGCTTTTCAGTTTTCGCTCGCGGAGTCTTATACGTTTGTGTGGAGCGTGGGAT 420
DB 769 GTGCTGCTTTTCAGTTTTCGCTCGCGGAGTCTTATACGTTTGTGTGGAGCGTGGGAT 828
QY 421 TCCAGTAATGACACCGGTTCAACCTGACAGTATTATTGAAAAGGCTTCTCACTCGGGCATG 480
DB 829 TCCAGTAATGACACCGGTTCAACCTGACAGTATTATTGAAAAGGCTTCTCACTCGGGCATG 888
QY 481 ATCAACCCAGCGCGGAGTGGCAGACGCTGAAGCAGAACACGGCGGTTCGCCACTTTGAG 540
DB 889 ATCAACCCAGCGCGGAGTGGCAGACGCTGAAGCAGAACACGGCGGTTCGCCACTTTGAG 948
QY 541 TATCAGATCCGCTGACCTGTGATGACTACTATATGGCTTTGGCTGTAATAGTTCTGCG 600
DB 949 TATCAGATCCGCTGACCTGTGATGACTACTATATGGCTTTGGCTGTAATAGTTCTGCG 1008
QY 601 CGCCCCAGAGATGACTTCTTTGGACACTATGCTGTGACCCAGAAATGGCAACAAACCTTGC 660
DB 1009 CGCCCCAGAGATGACTTCTTTGGACACTATGCTGTGACCCAGAAATGGCAACAAACCTTGC 1068
QY 661 ATGGAAGGCTGGATGGGCCCGGAAATGTAACAGAGCTATTTCGCCGACAGGCTGAGTCTCT 720
DB 1069 ATGGAAGGCTGGATGGGCCCGGAAATGTAACAGAGCTATTTCGCCGACAGGCTGAGTCTCT 1128
QY 721 AAGCATGGGTCTTGCACAACTCCAGAGTACTGAGGTGCGAGTAYGGCTGCGCAAGGCTG 780
DB 1129 AAGCATGGGTCTTGCACAACTCCAGAGTACTGAGGTGCGAGTAYGGCTGCGCAAGGCTG 1188
QY 781 TACTGTGATAAGTGCATCCACACCCGGGATGCGTCCACGGCATCTGTAAATGAGCCCTGG 840
DB 1189 TACTGTGATAAGTGCATCCACACCCGGGATGCGTCCACGGCATCTGTAAATGAGCCCTGG 1248
QY 841 CAGTGTCTGTGAGACCAACTGGGGCGGCAGCTCTGTGACAAAGATCTCAATTAATCTGT 900
DB 1249 CAGTGTCTGTGAGACCAACTGGGGCGGCAGCTCTGTGACAAAGATCTCAATTAATCTGT 1308
QY 901 GGGACTCATCAGCGCTGTCTCAACGGGGAACTTTGTAGCAACACAGGCCCTTGACAAATAT 960
DB 1309 GGGACTCATCAGCGCTGTCTCAACGGGGAACTTTGTAGCAACACAGGCCCTTGACAAATAT 1368
QY 961 CAGTGTCTGTGCTGAGGGGTTATCAGGACCCAACTCTGAAATTTGCTGAGCAGCGCTGC 1020
DB 1369 CAGTGTCTGTGCTGAGGGGTTATCAGGACCCAACTCTGAAATTTGCTGAGCAGCGCTGC 1428
QY 1021 CTCTGTATCCCTGTCAACAGAGCGAGCTGTAAAGAGACCTCCCTGGGCTTTGAGTGT 1080
DB 1429 CTCTGTATCCCTGTCAACAGAGCGAGCTGTAAAGAGACCTCCCTGGGCTTTGAGTGT 1488
QY 1081 GAGTGTTCCTCCAGGCTGGACCGGCCCCACATGCTCTACAAAATTTGATGACTGTTCTCCT 1140
DB 1489 GAGTGTTCCTCCAGGCTGGACCGGCCCCACATGCTCTACAAAATTTGATGACTGTTCTCCT 1548
QY 1141 AATAACTGTTCACACGGGGGACCTTCGCCAGGACCTGGTTAAACGGATTTAAGTGTGTGC 1200
DB 1549 AATAACTGTTCACACGGGGGACCTTCGCCAGGACCTGGTTAAACGGATTTAAGTGTGTGC 1608
QY 1201 CCCCCACAGTGGACTGGGAAAACGTCGCCAGTTAGATGCAATGAAATGTGAGGCGCAACCT 1260
DB 1609 CCCCCACAGTGGACTGGGAAAACGTCGCCAGTTAGATGCAATGAAATGTGAGGCGCAACCT 1668
QY 1261 TGTGTAACCGCAAAATCTCTGTAGAAATCTCATTCGCCAGCTTACTTGCAGCTGTTCTCCC 1320
DB 1669 TGTGTAACCGCAAAATCTCTGTAGAAATCTCATTCGCCAGCTTACTTGCAGCTGTTCTCCC 1728
QY 1321 GGCTGATGGGTGAGAAATTTGACATAAAATTTAATGACTGCTTGGCCAGTGTGAGT 1380
DB 1729 GGCTGATGGGTGAGAAATTTGACATAAAATTTAATGACTGCTTGGCCAGTGTGAGT 1788
QY 1381 GAGCCTCTCTGTGGGATTTGGTTAATGGTTATCGCTGTATCTGTGCTCACTGGCTATGCA 1440

Db 1789 GAGCCCTCTGTGGGATTTGGTTAATGGTTATCGCTGTATCTGTCCACCTGGCTATGCA 1848
Qy 1441 GGGGATCACTGTGAGAGAGACATCGATGAATGTGCCAGCAACCCCTGTGTTGAATGGGGT 1500
Db 1849 GGGGATCACTGTGAGAGAGACATCGATGAATGTGCCAGCAACCCCTGTGTTGAATGGGGT 1908
Qy 1501 CACTGTGAGAAATCAACAGATTCCAGTGTCTGTGTCCCACTGGTTTCTCTGGAAC 1560
Db 1909 CACTGTGAGAAATCAACAGATTCCAGTGTCTGTGTCCCACTGGTTTCTCTGGAAC 1968
Qy 1561 CTCTGTGAGTGGACATCGATTATTTGTGAGCTTAATCCCTGCCAGAAACGGTGGCCAGTGC 1620
Db 1969 CTCTGTGAGTGGACATCGATTATTTGTGAGCTTAATCCCTGCCAGAAACGGTGGCCAGTGC 2028
Qy 1621 TACAAACCGTGCAGTGAATTTCTGCAAGTGTGCCAGCACTATGAGGGCAGAACTGC 1680
Db 2029 TACAAACCGTGCAGTGAATTTCTGCAAGTGTGCCAGCACTATGAGGGCAGAACTGC 2088
Qy 1681 TCACACCTGAAAGACCACTGCGGCAGCACCCCTGTGAAAGTATTGACAGCTGCACAGTG 1740
Db 2089 TCACACCTGAAAGACCACTGCGGCAGCACCCCTGTGAAAGTATTGACAGCTGCACAGTG 2148
Qy 1741 GCCATGGCTTCCAAACGACACACCTGAAAGGGGTGGGTATATTTCTCCAAACGCTGTGGT 1800
Db 2149 GCCATGGCTTCCAAACGACACACCTGAAAGGGGTGGGTATATTTCTCCAAACGCTGTGGT 2208
Qy 1801 CCTCAGCGGAAGTCCAGAGTCACTCGGGAGGGGGAATTCACCTGTGACTGTGTAACAAAGGC 1860
Db 2209 CCTCAGCGGAAGTCCAGAGTCACTCGGGAGGGGGAATTCACCTGTGACTGTGTAACAAAGGC 2268
Qy 1861 TTCACGGGAACATACCTGCATGAAATATTAAGTGTGAGAGCAACCCCTGTGAGAAAC 1920
Db 2269 TTCACGGGAACATACCTGCATGAAATATTAAGTGTGAGAGCAACCCCTGTGAGAAAC 2328
Qy 1921 GGTGGCACTTGCATCGATGGTCAACTCTTACAGTGCATCTGTAGTGCAGCGTGGGAG 1980
Db 2329 GGTGGCACTTGCATCGATGGTCAACTCTTACAGTGCATCTGTAGTGCAGCGTGGGAG 2388
Qy 1981 GGGGCGCTACTGTGAAACCAATTAATGA CTGAGCCAGCAACCCCTGCGCAATGGGGGC 2040
Db 2389 GGGGCGCTACTGTGAAACCAATTAATGA CTGAGCCAGCAACCCCTGCGCAATGGGGGC 2448
Qy 2041 ACGTGTGCGACCTGGTCAAGTCTTACTGTGCACTGTAAATGSGTGGAAAGGAAG 2100
Db 2449 ACGTGTGCGACCTGGTCAAGTCTTACTGTGCACTGTAAATGSGTGGAAAGGAAG 2508
Qy 2101 ACCTGCCACTCAGCTGACAGTGTGATGAGGCCACGTGCACAAACGGTGGCACCTGC 2160
Db 2509 ACCTGCCACTCAGCTGACAGTGTGATGAGGCCACGTGCACAAACGGTGGCACCTGC 2568
Qy 2161 TATGATGAGGGGATGCTTTTAAGTGCATGTCTGTGGCGCTGGGAAGGAACCTGT 2220
Db 2569 TATGATGAGGGGATGCTTTTAAGTGCATGTCTGTGGCGCTGGGAAGGAACCTGT 2628
Qy 2221 AACATAGCCCGAAGCAGTGTGCTGCCCAACCCCTGCCATAATGGGGCACAATGTG 2280
Db 2629 AACATAGCCCGAAGCAGTGTGCTGCCCAACCCCTGCCATAATGGGGCACAATGTG 2688
Qy 2281 GTCAACGGCGAGTCTTTTACGTGGCTGTGCAAGGAAGGCTGGGAGGGGCCCATCTGTGCT 2340
Db 2689 GTCAACGGCGAGTCTTTTACGTGGCTGTGCAAGGAAGGCTGGGAGGGGCCCATCTGTGCT 2748
Qy 2341 CAGAAATCAATGACTGAGCCCTCATCCCTGTGTTACAAACAGGGGCACTGTGTGGATGGA 2400
Db 2749 CAGAAATCAATGACTGAGCCCTCATCCCTGTGTTACAAACAGGGGCACTGTGTGGATGGA 2808
Qy 2401 GACAACTGGTACCGGTGGAATGTGCCCGGGTTTGTGCGGCCGCACTGCAGAAATAAC 2460
Db 2809 GACAACTGGTACCGGTGGAATGTGCCCGGGTTTGTGCGGCCGCACTGCAGAAATAAC 2868
Qy 2461 ATCAATGAATGCCAGTCTTCACTTGTGCTTTGTGAGCGGACCTGTGTGGATGAGATCAAT 2520
Db 2869 ATCAATGAATGCCAGTCTTCACTTGTGCTTTGTGAGCGGACCTGTGTGGATGAGATCAAT 2928

Qy 2521 GGCTACCGGTGTGTCTGCCCTCCAGGGCACAGTGGTCCAAAGTGCCAGGAAGTTTCAGGG 2580
Db 2929 GGCTACCGGTGTGTCTGCCCTCCAGGGCACAGTGGTCCAAAGTGCCAGGAAGTTTCAGGG 2988
Qy 2581 AGACCTTGCATCAACCATGGGAGTGTGATACAGATGGGGCCAAATGGGATGATGACTGT 2640
Db 2989 AGACCTTGCATCAACCATGGGAGTGTGATACAGATGGGGCCAAATGGGATGATGACTGT 3048
Qy 2641 AATACTGCGCAGTGCCTGAAATGGAGGATCGCTGCTCAAGGTCTGTGTGGCCCTCGA 2700
Db 3049 AATACTGCGCAGTGCCTGAAATGGAGGATCGCTGCTCAAGGTCTGTGTGGCCCTCGA 3108
Qy 2701 CCTTGGCTGTCTCCAAAGGGCACAGGAGTGGCCAGCGGGCAGAGCTGCATCCCATC 2760
Db 3109 CCTTGGCTGTCTCCAAAGGGCACAGGAGTGGCCAGCGGGCAGAGCTGCATCCCATC 3168
Qy 2761 CTGGACGACCAAGTCTTCTGTCCACCCCTGCACCTGGTGTGGGCGAGTGTGGTCTTCAGT 2820
Db 3169 CTGGACGACCAAGTCTTCTGTCCACCCCTGCACCTGGTGTGGGCGAGTGTGGTCTTCAGT 3228
Qy 2821 CTCACGCGGTGAAGCAAAAGTGCACCTCTGACTCTCTATTTACCAGGATTAACCTGTGGAAC 2880
Db 3229 CTCACGCGGTGAAGCAAAAGTGCACCTCTGACTCTCTATTTACCAGGATTAACCTGTGGAAC 3288
Qy 2881 ATCACATTTTACCTTTTAAAGGAGATGATGTACACAGGTCTTACTACGAGGACATTTGC 2940
Db 3289 ATCACATTTTACCTTTTAAAGGAGATGATGTACACAGGTCTTACTACGAGGACATTTGC 3348
Qy 2941 AGTGAATTTGAGGAATTTGAATATTTTGAAGAATTTTTCGGCTGAATTAATCAATCTCAATC 3000
Db 3349 AGTGAATTTGAGGAATTTGAATATTTTGAAGAAATTTTTCGGCTGAATTAATCAATCTCAATC 3408
Qy 3001 GCTTGGAGGCTTCCCTTTCAGCGAACAATGAATACATGTGGCCATTTCTGCTGAGAT 3060
Db 3409 GCTTGGAGGCTTCCCTTTCAGCGAACAATGAATACATGTGGCCATTTCTGCTGAGAT 3468
Qy 3061 ATACGGGATGATGGGAACCCGATCAAGGAAATCACTGACAAAAATTAATCTGTTAGT 3120
Db 3469 ATACGGGATGATGGGAACCCGATCAAGGAAATCACTGACAAAAATTAATCTGTTAGT 3528
Qy 3121 AAAAGTGTGATGAAACAGCTCGCTGATTTGCTGCCCTGCGAGAAAGTAAGAGTTCAGAGCGG 3180
Db 3529 AAAAGTGTGATGAAACAGCTCGCTGATTTGCTGCCCTGCGAGAAAGTAAGAGTTCAGAGCGG 3588
Qy 3181 CCTCTGAAGAACAGAAACAGATTTCTTGTTCCTTGTGAGCTCTGTCTTAACTGTGGCT 3240
Db 3589 CCTCTGAAGAACAGAAACAGATTTCTTGTTCCTTGTGAGCTCTGTCTTAACTGTGGCT 3648
Qy 3241 TGGATCTTGTGTGGTGAAGCGCTTCTACTGTGTGCTGCGGAAGCGCGGAAGCCGGGC 3300
Db 3649 TGGATCTTGTGTGGTGAAGCGCTTCTACTGTGTGCTGCGGAAGCGCGGAAGCCGGGC 3708
Qy 3301 AGCCACACACACTCAGCTCTGAGGACAAACCAACCAACCGTGGGGAGCAGCTGAAC 3360
Db 3709 AGCCACACACACTCAGCTCTGAGGACAAACCAACCAACCGTGGGGAGCAGCTGAAC 3768
Qy 3361 CAGATCAAAAAACCCATTGAGAAACATGGGGCCAAACGGTCCCCATCAAGGATTAAG 3420
Db 3769 CAGATCAAAAAACCCATTGAGAAACATGGGGCCAAACGGTCCCCATCAAGGATTAAG 3828
Qy 3421 AACAGAACTCCAAAAATGTCTTAAATAAGGACACAAATTTCTGAAGTAGAAGAGGACGAC 3480
Db 3829 AACAGAACTCCAAAAATGTCTTAAATAAGGACACAAATTTCTGAAGTAGAAGAGGACGAC 3888
Qy 3481 ATGGAACAAACCCAGCAGAAAGCCCGTGGCAGCAGCGCGGTATACGCTGGTAGAC 3540
Db 3889 ATGGAACAAACCCAGCAGAAAGCCCGTGGCAGCAGCGCGGTATACGCTGGTAGAC 3948
Qy 3541 AGAAGAGAGAGCCCCCAACGGGACCGCCGACAAAACCCAACTGGAACAAACAG 3600
Db 3949 AGAAGAGAGAGCCCCCAACGGGACCGCCGACAAAACCCAACTGGAACAAACAG 4008

Qy	3601	GACAAACAGAGACTTTGGAAAGTCCCGACAGCTTAAACCGAATGGAGTACATCGTATAG	3657
Db	4009	GACAAACAGAGACTTTGGAAAGTCCCGACAGCTTAAACCGAATGGAGTACATCGTATAG	4065
RESULT 9			
US-09-855-722-6			
; Sequence 6, Application US/09855722			
; Patent No. 6638741			
; GENERAL INFORMATION:			
; APPLICANT: Sakano, Seiji			
; APPLICANT: Itoh, Akira			
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE			
; FILE REFERENCE: KP-8576			
; CURRENT APPLICATION NUMBER: US/09/855,722			
; CURRENT FILING DATE: 2001-05-16			
; PRIOR APPLICATION NUMBER: 09/214,278			
; PRIOR FILING DATE: 1999-01-26			
; NUMBER OF SEQ ID NOS: 32			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 6			
; LENGTH: 4208			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (409)..(4062)			
; NAME/KEY: sig_peptide			
; LOCATION: (409)..(501)			
; NAME/KEY: mat_peptide			
; LOCATION: (502)..(4062)			
US-09-855-722-6			
Query Match 99.8%; Score 3647.8; DB 3; Length 4208;			
Best Local Similarity 99.8%; Pred. No. 0;			
Matches 3649; Conservative 3; Mismatches 5; Indels 0; Gaps 0;			
Qy	1	ATGCGTTTCCCAACGGACRCGCGCGGTCTCGGGCGCCCTTAAGCCTCTCTGCTCGCCCTG	60
Db	409	ATGCGTTTCCCAACGGACRCGCGCGGTCTCGGGCGCCCTTAAGCCTCTCTGCTCGCCCTG	468
Qy	61	CTCTGTGTCCTTCGGAGCCAGAGTGTGTGGGGCCCTCGGGTCAGTTCGATTTGGAGATCCCTG	120
Db	469	CTCTGTGTCCTTCGGAGCCAGAGTGTGTGGGGCCCTCGGGTCAGTTCGATTTGGAGATCCCTG	528
Qy	121	TCCATGACAGAACTGTAACGGGAGCTGCAGAACGGGAACTGCTGCGCGCGCCCGCGAAC	180
Db	529	TCCATGACAGAACTGTAACGGGAGCTGCAGAACGGGAACTGCTGCGCGCGCCCGCGAAC	588
Qy	181	CCGGAGACCCGCAAGTGACACCCGCGAGGTGTGACATATCTCAAAGTGTGCTCAAG	240
Db	589	CCGGAGACCCGCAAGTGACACCCGCGAGGTGTGACATATCTCAAAGTGTGCTCAAG	648
Qy	241	GAGTATCAGTCCCGCGTACGGCCGGGGGGCCCTGAGCTTCGGCTCAGGGTCCAGCCT	300
Db	649	GAGTATCAGTCCCGCGTACGGCCGGGGGGCCCTGAGCTTCGGCTCAGGGTCCAGCCT	708
Qy	301	GTATCGGGGGCAACACTTTCAACCTCAAGCCAGCCGCGCAACGACCGCAACCCGATC	360
Db	709	GTATCGGGGGCAACACTTTCAACCTCAAGCCAGCCGCGCAACGACCGCAACCCGATC	768
Qy	361	GTGCTGCTTTTCAGTTTTCGCTCGCCGAGTCCCTATACGTTGTGTTGGAGGGCTGGGAT	420
Db	769	GTGCTGCTTTTCAGTTTTCGCTCGCCGAGTCCCTATACGTTGTGTTGGAGGGCTGGGAT	828
Qy	421	TCAGTAATGACACCGTTTCAACCTGACAGTATTATTGAAAGGCTTCTCACTCGGGCATG	480
Db	829	TCCAGTAATGACACCGTTTCAACCTGACAGTATTATTGAAAGGCTTCTCACTCGGGCATG	888
Qy	481	ATCAACCCAGCCGAGTGGCAGCGCTCAAGCAGAACACCGGGCGTTGCCACTTTGAG	540
Db	889	ATCAACCCAGCCGAGTGGCAGCGCTCAAGCAGAACACCGGGCGTTGCCACTTTGAG	948

|||||
Db 2029 TACACCGTCCAGTGACTATTTCTGCAAGTGCCTCCGAGGACTATGAGGGCAAGACTGC 2088
Qy 1681 TCACACTGAAAGACCACTGCGGACGACCCCTGTGGAAGTATGACAGCTGCACAGTG 1740
Db 2089 TCACACTGAAAGACCACTGCGGACGACCCCTGTGGAAGTATGACAGCTGCACAGTG 2148
Qy 1741 GCCATGGCTTCCAAACGACACACTGAGGGGTGGGTATATTTCCCTCCAAAGCTGTGGT 1800
Db 2149 GCCATGGCTTCCAAACGACACACTGAGGGGTGGGTATATTTCCCTCCAAAGCTGTGGT 2208
Qy 1801 CCTCACGGGAAGTGCAGAGTCACTGCGGAGGCAAAATTCACCTGTGACTGTATTAACAAAGGC 1860
Db 2209 CCTCACGGGAAGTGCAGAGTCACTGCGGAGGCAAAATTCACCTGTGACTGTATTAACAAAGGC 2268
Qy 1861 TTCAACGGGAACATCTGCCATGAAATATTAATGACTGTGAGAGCAACCCCTGTGAGAAAC 1920
Db 2269 TTCAACGGGAACATCTGCCATGAAATATTAATGACTGTGAGAGCAACCCCTGTGAGAAAC 2328
Qy 1921 GTTGGCACTTGCACTGATGTTCACTCTACAGTGCATCTGTAGTGAAGCTGGGAG 1980
Db 2329 GTTGGCACTTGCACTGATGTTCACTCTACAGTGCATCTGTAGTGAAGCTGGGAG 2388
Qy 1981 GGGGCTTACTGTGAACCAATATTAATGACTGACGCGAGAACCCCTGCCCAATGGGGC 2040
Db 2389 GGGGCTTACTGTGAACCAATATTAATGACTGACGCGAGAACCCCTGCCCAATGGGGC 2448
Qy 2041 ACCTGTCCGACCTGTGCAATGACTTCTACTGTGACTGTGTAATAATGGGTGAAAGGAAAG 2100
Db 2449 ACCTGTCCGACCTGTGCAATGACTTCTACTGTGACTGTGTAATAATGGGTGAAAGGAAAG 2508
Qy 2101 ACCTGCCACTGACGTGACGTGATGAGGCGCAGCTGCAACAGCTGGCACTGC 2160
Db 2509 ACCTGCCACTGACGTGACGTGATGAGGCGCAGCTGCAACAGCTGGCACTGC 2568
Qy 2161 TATGATCAGGGGATGCTTTTAAGTGCATGTGCTCGCGCTGGGAAGGAACAACTGT 2220
Db 2569 TATGATCAGGGGATGCTTTTAAGTGCATGTGCTCGCGCTGGGAAGGAACAACTGT 2628
Qy 2221 AACATAGCCGAAACAGTAGCTGCTGCCAACCCCTGCCAATATGGGGGCAATGTGTG 2280
Db 2629 AACATAGCCGAAACAGTAGCTGCTGCCAACCCCTGCCAATATGGGGGCAATGTGTG 2688
Qy 2281 GTCAACGGCGAGTCTTTACGTGCGTCTGCAAGGAAGCTGGGAGGGCCCATCTGTGCT 2340
Db 2689 GTCAACGGCGAGTCTTTACGTGCGTCTGCAAGGAAGCTGGGAGGGCCCATCTGTGCT 2748
Qy 2341 CAGAAATACCAATGACTGCAGCCCTCATCCCTGTTTACAAACAGCGGCACCTGTGTGGATGGA 2400
Db 2749 CAGAAATACCAATGACTGCAGCCCTCATCCCTGTTTACAAACAGCGGCACCTGTGTGGATGGA 2808
Qy 2401 GACAACTGGTACCGGTGCGAATGTGCCCGGGTTTTGCTGGGCCCGACTGCAGAAATAAAC 2460
Db 2809 GACAACTGGTACCGGTGCGAATGTGCCCGGGTTTTGCTGGGCCCGACTGCAGAAATAAAC 2868
Qy 2461 ATCAATGAATCCAGTCTTACCTTGTGCTTTGAGGCGACTGTGTGGATGAGATCAAT 2520
Db 2869 ATCAATGAATCCAGTCTTCACTTGTGCTTTGAGGCGACTGTGTGGATGAGATCAAT 2928
Qy 2521 GGCTACCGGTGTGTCTGCCCTCCAGGGGACAGTGGTGCAGTGCAGGAAGTTTCAGGG 2580
Db 2929 GGCTACCGGTGTGTCTGCCCTCCAGGGGACAGTGGTGCAGTGCAGGAAGTTTCAGGG 2988
Qy 2581 AGACCTTGCAATCAACATGGGAGTGTGATACAGATGGGGCCAAATGGGATGATGACTGT 2640
Db 2989 AGACCTTGCAATCAACATGGGAGTGTGATACAGATGGGGCCAAATGGGATGATGACTGT 3048
Qy 2641 AATACCTGCCAGTGCCTGAATGGACGATCGCTGCTCAAGGCTGTGTGGGCCCTCGA 2700
Db 3049 AATACCTGCCAGTGCCTGAATGGACGATCGCTGCTCAAGGCTGTGTGGGCCCTCGA 3108
Qy 2701 CCTTGCCTGCTCCAAAGGGGACAGCGAGTGGCCCGGCGGAGCTGCATCCCCATC 2760
|||||

Db 3109 CCTTGCCTGCTCCACAAAGGGGACAGCGAGTGCCTCCAGCGGGCAGAGCTGCATCCCCATC 3168
Qy 2761 CTGGGACGACAGTGTCTTCCATCCCTCCACTGGTGTGGGAGTGTGGTCTTCCAGT 2820
Db 3169 CTGGGACGACAGTGTCTTCCATCCCTCCACTGGTGTGGGAGTGTGGTCTTCCAGT 3228
Qy 2821 CTCCAGCCGGTGAAGACAAAGTGCACTCTGACTCTCTATTTACCAGGATTAATGTGCGAAC 2880
Db 3229 CTCCAGCCGGTGAAGACAAAGTGCACTCTGACTCTCTATTTACCAGGATTAATGTGCGAAC 3288
Qy 2881 ATCACATTTACCTTTAAACAGGAGATGATGCACAGGTCTTACTACGAGGACATTTGC 2940
Db 3289 ATCACATTTACCTTTAAACAGGAGATGATGCACAGGTCTTACTACGAGGACATTTGC 3348
Qy 2941 AGTGAATTCAGGAAATTTGAATATTTTGAAGAATGTTTCCGCTGAAATTTCAATCTACATC 3000
Db 3349 AGTGAATTCAGGAAATTTGAATATTTTGAAGAATGTTTCCGCTGAAATTTCAATCTACATC 3408
Qy 3001 GCTTGGGAGCTTCCCTTTCAGCGAAACAATGAATAACATGTGGCCATTTCTGTGTAAGAT 3060
Db 3409 GCTTGGGAGCTTCCCTTTCAGCGAAACAATGAATAACATGTGGCCATTTCTGTGTAAGAT 3468
Qy 3061 ATACGGGATGATGGGAAACCCGATCAAGGAAATCACTGAACAAATAATTCGATCTTGTAGT 3120
Db 3469 ATACGGGATGATGGGAAACCCGATCAAGGAAATCACTGAACAAATAATTCGATCTTGTAGT 3528
Qy 3121 AAACGTGATGGAACACAGCTCGCTGATTGCTGCGTTGCAAGTAAGAGTTCAGAGCGG 3180
Db 3529 AAACGTGATGGAACACAGCTCGCTGATTGCTGCGTTGCAAGTAAGAGTTCAGAGCGG 3588
Qy 3181 CCTCTGAAGAACAGAACAGATTTCTTGTTCCTTCTGAGCTCTCTCTTAACTGTGGCT 3240
Db 3589 CCTCTGAAGAACAGAACAGATTTCTTGTTCCTTCTGAGCTCTCTCTTAACTGTGGCT 3648
Qy 3241 TGGATCTGTGTGTGTGACGCGCTTCTACTGTGTGCTGCGGAAGCGGCGGAAAGCGGGC 3300
Db 3649 TGGATCTGTGTGTGTGACGCGCTTCTACTGTGTGCTGCGGAAGCGGCGGAAAGCGGGC 3708
Qy 3301 AGCCACACACACTCAGCCTCTGAGGACACACACCAACCAACCTGCGGGAGCAGCTGAAC 3360
Db 3709 AGCCACACACACTCAGCCTCTGAGGACACACACCAACCAACCTGCGGGAGCAGCTGAAC 3768
Qy 3361 CAGATCAAAAAACCCCATTTGAGAAACATTTGGGGCCCAACACCGTCCCATCAAGGATTAAGAG 3420
Db 3769 CAGATCAAAAAACCCCATTTGAGAAACATTTGGGGCCCAACACCGTCCCATCAAGGATTAAGAG 3828
Qy 3421 AACAGAACTCCAAAATGTCTAAAATAAGGACACAAATTTCTGAAAGTAGAAGAGGACGAC 3480
Db 3829 AACAGAACTCCAAAATGTCTAAAATAAGGACACAAATTTCTGAAAGTAGAAGAGGACGAC 3888
Qy 3481 ATGGACAAACACACGACGAAAGCCCGTTTGGCAAGCAGCGCGGTATACGCTGGTAGAC 3540
Db 3889 ATGGACAAACACACGACGAAAGCCCGTTTGGCAAGCAGCGCGGTATACGCTGGTAGAC 3948
Qy 3541 AGAGAAGAGAGGCCCCCAACCGGACCGCCGACAAAACACCCAACTGGACAAAACAAACAG 3600
Db 3949 AGAGAAGAGAGGCCCCCAACCGGACCGCCGACAAAACACCCAACTGGACAAAACAAACAG 4008
Qy 3601 GACAAACAGAGACTTGGAAAGTGGCCAGAGCTTTAAACCGAATGGAGTACATCGTATAG 3657
Db 4009 GACAAACAGAGACTTGGAAAGTGGCCAGAGCTTTAAACCGAATGGAGTACATCGTATAG 4065
|||||

RESULT 10

US-08-400-159-5
; Sequence 5, Application US/08400159
; Patent No. 5869282
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowicz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.

Query Match	99.7%	Score 3646.2;	DB 3;	Length 6464;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 3648; Conservative	3;	Mismatches 6;	Indels 0;	Gaps 0;

Qy	1	ATCGGTTCCCAACGGACRCGGGGCGCGGTTCGGGGCGGCCCTTAAGCCTTCTGCTCGCCCTG	60
Db	371	ATCGGTTCCCAACGGACAACCGGGCGCGGTTCGGGGCGGCCCTTAAGCCTTCTGCTCGCCCTG	430
Qy	61	CTCTGTGCCTTGCAGGCAAAAGTGTTGGGGCCCTCGGGTCAGTTCGAGTTGGAGATCCTG	120
Db	431	CTCTGTGCCTTGCAGGCCAAGTGTTGGGGCCCTCGGGTCAGTTCGAGTTGGAGATCCTG	490
Qy	121	TCCATGAGAAACGTGAACCGGGAGCTGCAGAACGGGAACTGCTCGGGCGGCCCGCGAAC	180
Db	491	TCCATGAGAAACGTGAACCGGGAGCTGCAGAACCGGAACTGCTCGGGCGGCCCGCGAAC	550
Qy	181	CCGGGACACCGCAAGTGCACCCGCGACGAGTGTGACACATACTTCAAAGTGTGCTCTCAAG	240
Db	551	CCGGGACACCGCAAGTGCACCCGCGACGAGTGTGACACATACTTCAAAGTGTGCTCTCAAG	610
Qy	241	GAGTATCAGTCCCGCGTCACGGCCGGGGGGCCCTGCGAGCTTCGGCTCAGGGTCCACGCCT	300

QY	1381	GAGCCTCCTGTGGGATTTGGTTAATGGTTATCGCTGTATCTGTCCACCTGGCTATGCA	1440
DB	1751	GACGCTCCTGTGGGATTTGGTTAATGGTTATCGCTGTATCTGTCCACCTGGCTATGCA	1810
QY	1441	GGGATCACCTGTGAGAGACATCGATGAATGTGCCAGCAACCCCTGTGTTTGAATGGGGT	1500
DB	1811	GGGATCACCTGTGAGAGACATCGATGAATGTGCCAGCAACCCCTGTGTTTGAATGGGGT	1870
QY	1501	CATGTGAGAAATGAATCAACAGATTCAGTGTCTGTGTCCCACTGTGTTTCTCTGGAAC	1560
DB	1871	CATGTGAGAAATGAATCAACAGATTCAGTGTCTGTGTCCCACTGTGTTTCTCTGGAAC	1930
QY	1561	CTCTGTGAGTGAATCAACAGATTCAGTGTCTGTGTCCCACTGTGTTTCTCTGGAAC	1620
DB	1931	CTCTGTGAGTGAATCAACAGATTCAGTGTCTGTGTCCCACTGTGTTTCTCTGGAAC	1990
QY	1621	TACAACCGTGCAGTGAATCAACAGATTCAGTGTCTGTGTCCCACTGTGTTTCTCTGGAAC	1680
DB	1991	TACAACCGTGCAGTGAATCAACAGATTCAGTGTCTGTGTCCCACTGTGTTTCTCTGGAAC	2050
QY	1681	TCACACTGAAAGACCACTGCGGACAGACCCCTGTGAGTGAATGTGACAGTGCACAGTG	1740
DB	2051	TCACACTGAAAGACCACTGCGGACAGACCCCTGTGAGTGAATGTGACAGTGCACAGTG	2110
QY	1741	GCCATGGCTTCCAAACGACACACCTGAAGGGGTGGGTATATTTCTTCCCAACGCTGTGGT	1800
DB	2111	GCCATGGCTTCCAAACGACACACCTGAAGGGGTGGGTATATTTCTTCCCAACGCTGTGGT	2170
QY	1801	CCTCAGCGGAAGTCAAGAGTCACTGCGGAGGCAAAATTCACCTGTGACTGTGAACAAAGGC	1860
DB	2171	CCTCAGCGGAAGTCAAGAGTCACTGCGGAGGCAAAATTCACCTGTGACTGTGAACAAAGGC	2230
QY	1861	TTCAACGGGAACATCTGCCATGAATAATTAATGACTGTGAGAGCAACCCCTGTGAGAAC	1920
DB	2231	TTCAACGGGAACATCTGCCATGAATAATTAATGACTGTGAGAGCAACCCCTGTGAGAAC	2290
QY	1921	GGTGGCACTTGATCGATGTGTCACCTCTACAGTGCATCTGTAGTGCAGCTGGAG	1980
DB	2291	GGTGGCACTTGATCGATGTGTCACCTCTACAGTGCATCTGTAGTGCAGCTGGAG	2350
QY	1981	GGGCGCTACTGTGAAACCAATATTAATGACTGTGAGAGCAACCCCTGTGAGAAC	2040
DB	2351	GGGCGCTACTGTGAAACCAATATTAATGACTGTGAGAGCAACCCCTGTGAGAAC	2410
QY	2041	ACGTGTGCGGACCTGTGCAATGACTTCTAATGACTGTGAGAGCAACCCCTGTGAGAAC	2100
DB	2411	ACGTGTGCGGACCTGTGCAATGACTTCTAATGACTGTGAGAGCAACCCCTGTGAGAAC	2470
QY	2101	ACCTGCCACTCAGTGCAGTGCATGAGGCGCAGCTGCAACACGGTGGCACCTGC	2160
DB	2471	ACCTGCCACTCAGTGCAGTGCATGAGGCGCAGCTGCAACACGGTGGCACCTGC	2530
QY	2161	TATGATGAGGGGATGCTTTTAAAGTGCATGTCTGCGGCTGGGAGGAAACCACTGT	2220
DB	2531	TATGATGAGGGGATGCTTTTAAAGTGCATGTCTGCGGCTGGGAGGAAACCACTGT	2590
QY	2221	AACATAGCCGAAACAGTGTGCTGCCAACCCCTGCCATATAGGGGCGACATGTGTG	2280
DB	2591	AACATAGCCGAAACAGTGTGCTGCCAACCCCTGCCATATAGGGGCGACATGTGTG	2650
QY	2281	GTCAACGGGAGTCTTTTAAAGTGCATGTCTGCGGCTGGGAGGAAACCACTGTGCT	2340
DB	2651	GTCAACGGGAGTCTTTTAAAGTGCATGTCTGCGGCTGGGAGGAAACCACTGTGCT	2710
QY	2341	CAGATACCAATGACTGCAGGCTCATCTCTGTTTACAAACAGCGGCACCTGTGTGATGGA	2400
DB	2711	CAGATACCAATGACTGCAGGCTCATCTCTGTTTACAAACAGCGGCACCTGTGTGATGGA	2770
QY	2401	GACAACTGGTACCGGTGCAATGTGCCCCGGGTTTGTGGGCCCCGACTGCAGAAATAAAC	2460
DB	2771	GACAACTGGTACCGGTGCAATGTGCCCCGGGTTTGTGGGCCCCGACTGCAGAAATAAAC	2830

QY	2461	ATCAATGAATGCCAGTCTTTCACCTTGTGCTTTGGAGCGACCTGTGTGGATGAGATCAAT	2520
DB	2831	ATCAATGAATGCCAGTCTTTCACCTTGTGCTTTGGAGCGACCTGTGTGGATGAGATCAAT	2890
QY	2521	GGCTACCGGTGTGTCTGCCCTCCAGGGCAAGTGTGTGCCAAAGTGCAGGAAGTTTCAGGG	2580
DB	2891	GGCTACCGGTGTGTCTGCCCTCCAGGGCAAGTGTGTGCCAAAGTGCAGGAAGTTTCAGGG	2950
QY	2581	AGACCTTGTGATCACCAATGGGGAGTGTATACAGATGGGGGCCAAATGGGGATGATGACTGT	2640
DB	2951	AGACCTTGTGATCACCAATGGGGAGTGTATACAGATGGGGGCCAAATGGGGATGATGACTGT	3010
QY	2641	AATACCTGCGCAGTCCCTGAATGGAGCGGATCGCTGTCTCAAAGCTCTGGTGTGGCCCTCGA	2700
DB	3011	AATACCTGCGCAGTCCCTGAATGGAGCGGATCGCTGTCTCAAAGCTCTGGTGTGGCCCTCGA	3070
QY	2701	CCTTGTGCTGCTCCAAAAGGGCAACGAGTGTGCCAGGGGCGAGAGTGCATCCCCATC	2760
DB	3071	CCTTGTGCTGCTCCAAAAGGGCAACGAGTGTGCCAGGGGCGAGAGTGCATCCCCATC	3130
QY	2761	CTGGACGACAGTGTCTGCTCCACCCCTGCACTGTGTGTGGCGAGTGTGGTCTTCCAGT	2820
DB	3131	CTGGACGACAGTGTCTGCTCCACCCCTGCACTGTGTGTGGCGAGTGTGGTCTTCCAGT	3190
QY	2821	CTCCAGCCGTGAAAGCAAAAGTGCACCTCTGACTCTCTATTTACCCAGGATAACTGTGCGAAC	2880
DB	3191	CTCCAGCCGTGAAAGCAAAAGTGCACCTCTGACTCTCTATTTACCCAGGATAACTGTGCGAAC	3250
QY	2881	ATCACAATTTTACCTTTAAACAGGAGATGTGTCAACAGGTCTTACTACGAGCAATTTTGTG	2940
DB	3251	ATCACAATTTTACCTTTAAACAGGAGATGTGTCAACAGGTCTTACTACGAGCAATTTTGTG	3310
QY	2941	AGTGAATTTGAGGAATTTGAATATTTTGAAGAAATTTTCCGCTGAATATTTCAATCTACATC	3000
DB	3311	AGTGAATTTGAGGAATTTGAATATTTTGAAGAAATTTTCCGCTGAATATTTCAATCTACATC	3370
QY	3001	GCTTGGAGGCTTCCCTTCCAGGAAACAAATGAAATACATGTGGCCCAATTTCTGTGGAAGAT	3060
DB	3371	GCTTGGAGGCTTCCCTTCCAGGAAACAAATGAAATACATGTGGCCCAATTTCTGTGGAAGAT	3430
QY	3061	ATAACGGATGATGGGAACCCGATCAAGGAAATCACTGACAAAATAATCTGATCTTTGTAGT	3120
DB	3431	ATAACGGATGATGGGAACCCGATCAAGGAAATCACTGACAAAATAATCTGATCTTTGTACT	3490
QY	3121	AAACGTGTGAAACAGCTCGCTGATTTCTGCGTTCGAGAACTAAGAGTTTCAGAGCGG	3180
DB	3491	AAACGTGTGAAACAGCTCGCTGATTTCTGCGTTCGAGAACTAAGAGTTTCAGAGCGG	3550
QY	3181	CCTCTGAAGAAACAGAAACAGATTTTCTTGTTCCTTGTCTGAGCTCTGTCTTAACTGTGGCT	3240
DB	3551	CCTCTGAAGAAACAGAAACAGATTTTCTTGTTCCTTGTCTGAGCTCTGTCTTAACTGTGGCT	3610
QY	3241	TGGATCTGTGTGTGACGGCTTCTACTGTGTGCTGCGAAGCGGGCGAAGCCGGGC	3300
DB	3611	TGGATCTGTGTGTGACGGCTTCTACTGTGTGCTGCGAAGCGGGCGAAGCCGGGC	3670
QY	3301	AGCCACACACTCAGCCCTCTGAGGACAAACACCAACCAACCTGCGGGAGCGCTCAAC	3360
DB	3671	AGCCACACACTCAGCCCTCTGAGGACAAACACCAACCAACCTGCGGGAGCGCTCAAC	3730
QY	3361	CAGATCAAAAACCCCACTTGAAGAAACATGGGGCCAAACCGCTCCCATCAAGGATTAAGAG	3420
DB	3731	CAGATCAAAAACCCCACTTGAAGAAACATGGGGCCAAACCGCTCCCATCAAGGATTAAGAG	3790
QY	3421	AACAAGAACTCCAAAATGTCTTAAATAAGGACACAAATTTCTGAAGTAGAAGGAGACGAC	3480
DB	3791	AACAAGAACTCCAAAATGTCTTAAATAAGGACACAAATTTCTGAAGTAGAAGGAGACGAC	3850
QY	3481	ATGGAACAAACCCAGCAGAAAGCCGGTTTGGCAAGCGCGGTATACCTGCTGTTAGAC	3540
DB	3851	ATGGAACAAACCCAGCAGAAAGCCGGTTTGGCAAGCGCGGTATACCTGCTGTTAGAC	3910
QY	3541	AGAGAAGAGAAGCCCCCAACCGGCAACCGGCAACCAACCAACCTGGACCAACAAACAG	3600

431	CTCTGTGCCCTCGAGCGCAAGGTGTGTGGGCGCTCTCGGGTCAGTTCGAGTTGGAGATCCTG	490
121	TCCATGACGAAACGTGAACGGGGAGCTGCGAAACGGGAACCTGCTCGCGCGCGCCCGGAAC	180
491	TCCAATGACGAAACGTGAACGGGGAGCTGCGAAACGGGAACCTGCTCGCGCGCGCCCGGAAC	550
181	CCGGGACACCGCAAGTGTCACCCGGGACGAGTGTGACACATATCTTCAAAGTGTGCCTCAAG	240
551	CCGGGACACCGCAAGTGTCACCCCGCGACGAGTGTGACACATATCTTCAAAGTGTGCCTCAAG	610
241	GAGTATCAGTCCCGCGTTCAGCGCCGGGGGGCCCTGACAGCTTCGGCTCAGGGTCCACGCCT	300
611	GAGTATCAGTCCCGCGTTCAGCGCCGGGGGGCCCTGACAGCTTCGGCTCAGGGTCCACGCCT	670
301	GTCAATCGGGGGCAACACCTTCAACCTCAAGGCCAGCGCGGCAACGACCGCAACCGCATC	360
671	GTCAATCGGGGGCAACACCTTCAACCTCAAGGCCAGCGCGGCAACGACCGCATC	730
361	GTGCTGCTTTTCAAGTTTCGGCTGGCGCAGGTCTTAATACGTTGCTTGTGGAGGCGTGGGAT	420
731	GTGCTGCTTTTCAAGTTTCGGCTGGCGCAGGTCTTAATACGTTGCTTGTGGAGGCGTGGGAT	790
421	TCCAGTAATACACACCGTTCAACCTGACAGTATTAATTGAAAAGGCTTCTCACTCGGGCATG	480
791	TCCAGTAATACACACCGTTCAACCTGACAGTATTAATTGAAAAGGCTTCTCACTCGGGCATG	850
481	ATCAACCCCGAGCGGAGTGGCAGACGCTGAAGCAGAAACACGGGCGTTGCCCACTTTGAG	540
851	ATCAACCCCGAGCGGAGTGGCAGACGCTGAAGCAGAAACACGGGCGTTGCCCACTTTGAG	910
541	TATCAGATCCGGGTGACCTGTGTATGACTACTATATGGCTTTTGGCTGYAATAAGTTCTGC	600
911	TATCAGATCCGGGTGACCTGTGTATGACTACTATATGGCTTTTGGCTGYAATAAGTTCTGC	970
601	CGCCCCAGAGATGACTTCTTTTGGACACTATGCTGTGACCCAGAGTGCACAAACCTTGC	660
971	CGCCCCAGAGATGACTTCTTTTGGACACTATGCTGTGACCCAGAGTGCACAAACCTTGC	1030
661	ATGGAAGGCTGGATGGGCCCCGAATGTAAACAGAGCTATTTGCCGCAAGGCTGCAGTCTCT	720
1031	ATGGAAGGCTGGATGGGCCCCGAATGTAAACAGAGCTATTTGCCGCAAGGCTGCAGTCTCT	1090
721	AAGCATGGGTCTTGCAAACTCCCAGGTGACTGCAGGTGCCAGTAYGGCTGGCAAGGCGTG	780
1091	AAGCATGGGTCTTGCAAACTCCCAGGTGACTGCAGGTGCCAGTAYGGCTGGCAAGGCGTG	1150
781	TACTGTGATAAGTGCATCCACACCCGGGATCGGTCCACGGCATCTGTATAGACCCCTGG	840
1151	TACTGTGATAAGTGCATCCACACCCGGGATCGGTCCACGGCATCTGTATAGACCCCTGG	1210
841	CAGTGCCTCTGTGAGACCAACTGGGGCGGCGAGCTGTGACAAAGATCTCAATTACTGT	900
1211	CAGTGCCTCTGTGAGACCAACTGGGGCGGCGAGCTGTGACAAAGATCTCAATTACTGT	1270
901	GGGACTCATCAGCCGTGTCTCAACGGGGGAACTTGTAGCAACACAGGCCCTGACAAATAT	960
1271	GGGACTCATCAGCCGTGTCTCAACGGGGGAACTTGTAGCAACACAGGCCCTGACAAATAT	1330
961	CAGTGTTCCTGCCCTGAGGGGTATTACGACCACCACTGTGCAAAATTCGTAGACACGCCTGC	1020
1331	CAGTGTTCCTGCCCTGAGGGGTATTACGACCACCACTGTGCAAAATTCGTAGACACGCCTGC	1390
1021	CTCTCTGATCCCTGTGCACACAGAGGCAGCTGTAAAGGAGACCTCCCTGGGCTTTGAGTGT	1080
1391	CTCTCTGATCCCTGTGCACACAGAGGCAGCTGTAAAGGAGACCTCCCTGGGCTTTGAGTGT	1450
1081	GAGTGTTCCTCCAGGCTGGACCGGCCCCACATGCTCTACAAACATGTGATGACTGTTCTCCT	1140
1451	GAGTGTTCCTCCAGGCTGGACCGGCCCCACATGCTCTACAAACATGTGATGACTGTTCTCCT	1510
1141	AATAACTGTTCCACGGGGGCACTGCCAGGACCTGGTTAAACGGAATTAAGTGTGTGTGC	1200
1511	AATAACTGTTCCACGGGGGCACTGCCAGGACCTGGTTAAACGGAATTAAGTGTGTGTGC	1570

551	Db	CCGGGAGACCGCAAGTGCACCCGGACGAGTGTGACATAC TTTCAAGTGTGCCTCAAG	610
241	Qy	GAGTATCAGTCCCGCGTCAAGCGCGGGGGCCCTGCAGCTTCGGCTCAGGGTCCACGCCT	300
611	Db	GAGTATCAGTCCCGCGTCAAGCGGGGGCCCTGCAGCTTCGGCTCAGGGTCCACGCCT	670
301	Qy	GTATTCGGGGCAACACCTTCAACCTCAAGGCCAGCCGGCGCAACGACCCGCAACCGCATC	360
671	Db	GTATTCGGGGCAACACCTTCAACCTCAAGGCCAGCCGGCGCAACGACCCGCAACCGCATC	730
361	Qy	GTGCTGCTTTTCAGTTTCGGCTCGCGGAGGTCTTATACGTTCGTTGTGGAGGGGTGGAT	420
731	Db	GTGCTGCTTTTCAGTTTCGGCTCGCGGAGGTCTTATACGTTCGTTGTGGAGGGGTGGAT	790
421	Qy	TCCAGTAATGACACCGTTCAACCTGACAGTATTAATTGAAAAGGCTTCTCACTCGGGCATG	480
791	Db	TCCAGTAATGACACCGTTCAACCTGACAGTATTAATTGAAAAGGCTTCTCACTCGGGCATG	850
481	Qy	ATCAACCCAGCGCGGAGTGGGAGACGCTGAAGCAGAAACA CGGGCGTTTGCCCATTTTGAG	540
851	Db	ATCAACCCAGCGCGGAGTGGGAGACGCTGAAGCAGAAACA CGGGCGTTTGCCCATTTTGAG	910
541	Qy	TATCAGATCCGCTGACCTGTAATGACTACTACTATATGGCTTTGSGCTYAAATAGTTCCTGC	600
911	Db	TATCAGATCCGCTGACCTGTAATGACTACTACTATATGGCTTTGSGCTYAAATAGTTCCTGC	970
601	Qy	CGCCCCAGAGATGACTTTCTTTTGGACACTATATGCTGTGAC CAGAAATGSCAACAAACTTGC	660
971	Db	CGCCCCAGAGATGACTTTCTTTTGGACACTATATGCTGTGAC CAGAAATGSCAACAAACTTGC	1030
661	Qy	ATGGNAGGCTGGATGGGCCCCGGATGTAACAGAGCTATTTT GCCGACNAGGCTGCAGTCCCT	720
1031	Db	ATGGNAGGCTGGATGGGCCCCGGATGTAACAGAGCTATTTT GCCGACNAGGCTGCAGTCCCT	1090
721	Qy	AAGCATGGGTCTTTGCAAACTCCACAGGTGACTGCAGGTGCCAGTGCCAGTAYGGCTGGCAAGGCGCTG	780
1091	Db	AAGCATGGGTCTTTGCAAACTCCACAGGTGACTGCAGGTGCCAGTGCCAGTAYGGCTGGCAAGGCGCTG	1150
781	Qy	TACTGTGATAAGTGCAATCCACACCCGGGATCGTCCACGCACTGTGTAATAGACCCCTGG	840
1151	Db	TACTGTGATAAGTGCAATCCACACCCGGGATCGTCCACGCACTGTGTAATAGACCCCTGG	1210
841	Qy	CAGTGCTCTGTGAGACCAACTGGGGGGGCCAGCTCTGTGACAAAGNATCTCAATTACTGT	900
1211	Db	CAGTGCTCTGTGAGACCAACTGGGGGGGCCAGCTCTGTGACAAAGNATCTCAATTACTGT	1270
901	Qy	GGGACTCATCAGCCGTGTCTCAACGGGGGAACTTTGTAGCAACACAGGCCCTGACAAATAT	960
1271	Db	GGGACTCATCAGCCGTGTCTCAACGGGGGAACTTTGTAGCAACACAGGCCCTGACAAATAT	1330
961	Qy	CAGTGTTCCTGCCCTGAGGGGTATTCAGGACCCAACTGTGAAATTCGTGAGCAGCGCTGC	1020
1331	Db	CAGTGTTCCTGCCCTGAGGGGTATTCAGGACCCAACTGTGAAATTCGTGAGCAGCGCTGC	1390
1021	Qy	CTCTCTGATCCCTGTCAACACAGGAGCAGCTGTAAAGGAGACCTCCCTGGGCTTTGAGTGT	1080
1391	Db	CTCTCTGATCCCTGTCAACACAGGAGCAGCTGTAAAGGAGACCTCCCTGGGCTTTGAGTGT	1450
1081	Qy	GAGTGTTCCTCCAGGCTGGACCGGGCCCCACATGCTCTCAAAACATTTGATGACTGTTCTCCT	1140
1451	Db	GAGTGTTCCTCCAGGCTGGACCGGGCCCCACATGCTCTCAAAACATTTGATGACTGTTCTCCT	1510
1141	Qy	AATAACTGTTTCCACGGGGGCACTGTCAGGACCTGGTTTAA CGGATTTTAAAGTGTGTGTGC	1200
1511	Db	AATAACTGTTTCCACGGGGGCACTGTCAGGACCTGGTTTAA CGGATTTTAAAGTGTGTGTGC	1570

US-09-195-524-5					
; Sequence 5, Application US/09195524					
; Patent No. 6708499					
; GENERAL INFORMATION:					
; APPLICANT: Ish-Horowicz, David					
; APPLICANT: Henrique, Domingos M.P.					
; APPLICANT: Lewis, Julian H.					
; APPLICANT: Myat, Anna M.					
; APPLICANT: Fleming, Robert J.					
; APPLICANT: Artavanis-Tsakonas, Spyridon					
; APPLICANT: Mann, Robert S.					
; APPLICANT: Gray, Grace E.					
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE					
; TITLE OF INVENTION: SERATE GENE AND METHODS BASED THEREON					
; NUMBER OF SEQUENCES: 20					
; CORRESPONDENCE ADDRESS:					
; ADDRESSEE: Pennie & Edmonds					
; STREET: 1155 Avenue of the Americas					
; CITY: New York					
; STATE: New York					
; COUNTRY: U.S.A.					
; ZIP: 10036-2711					
; COMPUTER READABLE FORM:					
; MEDIUM TYPE: Floppy disk					
; COMPUTER: IBM PC compatible					
; OPERATING SYSTEM: PC-DOS/MS-DOS					
; SOFTWARE: Patent In Release #1.0, Version #1.30					
; CURRENT APPLICATION DATA:					
; APPLICATION NUMBER: US/09/195,524					
; FILING DATE:					
; CLASSIFICATION:					
; PRIOR APPLICATION DATA:					
; APPLICATION NUMBER: US 08/611,729					
; FILING DATE: 06-MAR-1996					
; ATTORNEY/AGENT INFORMATION: ;					
; NAME: Misrock, S. Leslie					
; REGISTRATION NUMBER: 18,872					
; REFERENCE/DOCKET NUMBER: 7326-037					
; TELECOMMUNICATION INFORMATION:					
; TELEPHONE: (212) 790-9090					
; TELEFAX: (212) 869-9741/8864					
; TELEX: 66141 PENNIE					
; INFORMATION FOR SEQ ID NO: 5:					
; SEQUENCE CHARACTERISTICS:					
; LENGTH: 6464 base pairs					
; TYPE: nucleic acid					
; STRANDEDNESS: double					
; TOPOLOGY: unknown					
; MOLECULE TYPE: cDNA					
; FEATURE:					
; NAME/KEY: CDS					
; LOCATION: 371..4024					
; US-09-195-524-5					
Query Match 99.7%; Score 3646.2; DB 3; Length 6464;					
Best Local Similarity 99.8%; Pred. No. 0;					
Matches 3648; Conservative 3; Mismatches 6; Indels 0; Gaps 0;					
QY	1	ATGCGTTCCTCCACGGACGCCGCGGTCCGGGCGCCCCCTTAAGCCTCCTGCCTCGCCCTG	60		
Dd	371	ATGGGTTCCTCCACGGACGCCGCGGTCCGGGCGCCCCCTTAAGCCTCCTGCCTG	430		
QY	61	CTCTGTGCCCTCGGACCAAGGTGTGTGGGCGCTGGGTCAAGTTTGAGTTGGAGATCCTG	120		

Qy	1201	CCCCCAGTGGACTGGGAAAACGTGCGAGTTAGATGCAAAATGAATGTGAGGCCAAACCT	1260
Db	1571	CCCCCAGTGGACTGGGAAAACGTGCGAGTTAGATGCAAAATGAATGTGAGGCCAAACCT	1630
Qy	1261	TGTGTAAACGCCAAATCCTGTGAAGAAATCTCATTTGCCAGCTACTACTCGACTGTCTTTCCC	1320
Db	1631	TGTGTAAACGCCAAATCCTGTGAAGAAATCTCATTTGCCAGCTACTACTCGACTGTCTTTCCC	1690
Qy	1321	GGCTGGATGGGTGAGAAATGTGACATATAATTAATGACTGCTTGGCCAGTGTGAGAAAT	1380
Db	1691	GGCTGGATGGGTGAGAAATGTGACATATAATTAATGACTGCTTGGCCAGTGTGAGAAAT	1750
Qy	1381	GACGCTCTCTGTGGGAAATTTGTTAAATGGTTATCGCTGTATCTGTCCACCTGGCTATGCA	1440
Db	1751	GACGCTCTCTGTGGGAAATTTGTTAAATGGTTATCGCTGTATCTGTCCACCTGGCTATGCA	1810
Qy	1441	GGCGATCACTGTGAGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGGT	1500
Db	1811	GGCGATCACTGTGAGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGGT	1870
Qy	1501	CACGTGCAGAAATGAAATCAACAGATTCAGTGTCTGTGTCCCACTGGTTTCTCTGGAAC	1560
Db	1871	CACGTGCAGAAATGAAATCAACAGATTCAGTGTCTGTGTCCCACTGGTTTCTCTGGAAC	1930
Qy	1561	CTCTGTGAGTGGACATCGATTTATGTGAGCCTAATCCCTGCCAGAACGGTGGCCAGTGC	1620
Db	1931	CTCTGTGAGTGGACATCGATTTATGTGAGCCTAATCCCTGCCAGAACGGTGGCCAGTGC	1990
Qy	1621	TACAAACGTCGACGTGACTATTTCTGCAAGTGGCCCCGAGGACTATGAGGGCAAGAACTGC	1680
Db	1991	TACAAACGTCGACGTGACTATTTCTGCAAGTGGCCCCGAGGACTATGAGGGCAAGAACTGC	2050
Qy	1681	TCACACTGAAAGACCACTGCGCGACAGACCCCTCTGTGAAAGTGAATGACGTGCACAGTG	1740
Db	2051	TCACACTGAAAGACCACTGCGCGACAGACCCCTCTGTGAAAGTGAATGACGTGCACAGTG	2110
Qy	1741	GCCATGCTTCCACGACACACCTGGAAGGGTGGGTATATTTCCCTCCACGCTCTGTGGT	1800
Db	2111	GCCATGCTTCCACGACACACCTGGAAGGGTGGGTATATTTCCCTCCACGCTCTGTGGT	2170
Qy	1801	CCTCAGCGGAAGTGCAGAGTCACTGCGGGAGGCAAAATTCACCTGTGACTGTGATGACAGTGC	1860
Db	2171	CCTCAGCGGAAGTGCAGAGTCACTGCGGGAGGCAAAATTCACCTGTGACTGTGATGACAGTGC	2230
Qy	1861	TTCAACGGGAACATCTGCGCATGAAATTAATGACTGTGTGAGAGCAACCCCTGTGAGAAAC	1920
Db	2231	TTCAACGGGAACATCTGCGCATGAAATTAATGACTGTGTGAGAGCAACCCCTGTGAGAAAC	2290
Qy	1921	GGTGGCACTTGCATCGATGGTGTCACTCCCTACAGTGCATCTGTAGTGGCTGGGAG	1980
Db	2291	GGTGGCACTTGCATCGATGGTGTCACTCCCTACAGTGCATCTGTAGTGGCTGGGAG	2350
Qy	1981	GGGGCTACTGTGAAACCAATTAATTAAGTACAGTGCAGCCAGCAACCCCTGCGCAAAATGGGGG	2040
Db	2351	GGGGCTACTGTGAAACCAATTAATTAAGTACAGTGCAGCCAGCAACCCCTGCGCAAAATGGGGG	2410
Qy	2041	ACGTGTGCGGACTGTGTAATGACTTCTTCTGTGACTGTGTAATAAATGGGTGAAAGGAAAG	2100
Db	2411	ACGTGTGCGGACTGTGTAATGACTTCTTCTGTGACTGTGTAATAAATGGGTGAAAGGAAAG	2470
Qy	2101	ACCTGGCACTCACTGACAGTCACTGTGATCAGGCAAGTGCACCAACGGTGGCACCTGC	2160
Db	2471	ACCTGGCACTCACTGACAGTCACTGTGATCAGGCAAGTGCACCAACGGTGGCACCTGC	2530
Qy	2161	TATGATGAGGGGATGCTTTTAAAGTGCATGTGTCTGTGGCGCTGGGAGGAAACACCTGT	2220
Db	2531	TATGATGAGGGGATGCTTTTAAAGTGCATGTGTCTGTGGCGCTGGGAGGAAACACCTGT	2590
Qy	2221	AACATAGCCCCGAAACAGTGTGCTGCTGCCAACCCCTGCGCAATATGGGGGACATGTGTG	2280
Db	2591	AACATAGCCCCGAAACAGTGTGCTGCTGCCAACCCCTGCGCAATATGGGGGACATGTGTG	2650

Qy	2281	GTCAACGGCGAGTCCTTTACGTGCGTCTGCAAGAAAGGTGGAGGGGCCCATCTGTGCT	2340
Db	2651	GTCAACGGCGAGTCCTTTACGTGCGTCTGCAAGAAAGGTGGAGGGGCCCATCTGTGCT	2710
Qy	2341	CAGAAATCAATGACTGTGAGCCCTCATCCCTGTTTACAAACAGCGGACCTGTGTGGATGGA	2400
Db	2711	CAGAAATCAATGACTGTGAGCCCTCATCCCTGTTTACAAACAGCGGACCTGTGTGGATGGA	2770
Qy	2401	GACAACTGGTACCGGTGGGAAATGTGCCCCGGGTTTTGTGGGCCCGACATGACAGAAATAAC	2460
Db	2771	GACAACTGGTACCGGTGGGAAATGTGCCCCGGGTTTTGTGGGCCCGACATGACAGAAATAAC	2830
Qy	2461	ATCAATGAATGCCAGTCTTTCACCTTGTGCTTTGGAGCGGACCTGTGTGGATGAGATCAAT	2520
Db	2831	ATCAATGAATGCCAGTCTTTCACCTTGTGCTTTGGAGCGGACCTGTGTGGATGAGATCAAT	2890
Qy	2521	GGCTACCGGTGTCTGCCCTCAGGGGCAAGTGGTGCCTCAAGTGCACAGGAAGTTTCAGGG	2580
Db	2891	GGCTACCGGTGTCTGCCCTCAGGGGCAAGTGGTGCCTCAAGTGCACAGGAAGTTTCAGGG	2950
Qy	2581	AGACCTTGCAATCAACATGGGGAGTGTATACAGATGGGGCCAAATGGGATGATGACTGT	2640
Db	2951	AGACCTTGCAATCAACATGGGGAGTGTATACAGATGGGGCCAAATGGGATGATGACTGT	3010
Qy	2641	AATACCTTGCAGTCCCTGAAATGAGACGGATCGCTCTGCTCAAAAGTCTGTGTGGCCCTCGA	2700
Db	3011	AATACCTTGCAGTCCCTGAAATGAGACGGATCGCTCTGCTCAAAAGTCTGTGTGGCCCTCGA	3070
Qy	2701	CCTTGTCTGTCTCAAAAGGGCAAGGAGTGGCCCCAGCGGGCAGAGCTGCATCCCCATC	2760
Db	3071	CCTTGTCTGTCTCAAAAGGGCAAGGAGTGGCCCCAGCGGGCAGAGCTGCATCCCCATC	3130
Qy	2761	CTGGAAGACAGTGTCTGCTCCACCTGCACCTGTGTGGCGGAGTGTGCTCTTCAGT	2820
Db	3131	CTGGAAGACAGTGTCTGCTCCACCTGCACCTGTGTGGCGGAGTGTGCTCTTCAGT	3190
Qy	2821	CTCAGCGGTGAAAGCAAAAGTGCACCTCTGACTCTCTATTATACAGGATAAATCTGTGGAAC	2880
Db	3191	CTCAGCGGTGAAAGCAAAAGTGCACCTCTGACTCTCTATTATACAGGATAAATCTGTGGAAC	3250
Qy	2881	ATCACATTTACCTTTTAAACAGGAGATGATGTGCACAGGTCTTACTACGGAGCATTGTGC	2940
Db	3251	ATCACATTTACCTTTTAAACAGGAGATGATGTGCACAGGTCTTACTACGGAGCATTGTGC	3310
Qy	2941	AGTGAATTTGAGGAATTTGAAATATTTTGAAGAAATTTTCCGCTGAATTAATCTACATC	3000
Db	3311	AGTGAATTTGAGGAATTTGAAATATTTTGAAGAAATTTTCCGCTGAATTAATCTACATC	3370
Qy	3001	GCTTGCAGGCTTTCCTTTCAGCGAAACAAATGAAATACATGTGGCCATTTCTGTGAAAGAT	3060
Db	3371	GCTTGCAGGCTTTCCTTTCAGCGAAACAAATGAAATACATGTGGCCATTTCTGTGAAAGAT	3430
Qy	3061	ATACGGGATGATGGGAACCCGATCAGGAAATCACTGACAAAATTAATCGATCTTTGTAGT	3120
Db	3431	ATACGGGATGATGGGAACCCGATCAGGAAATCACTGACAAAATTAATCGATCTTTGTACT	3490
Qy	3121	AAACGTGATGGAACACGCTCGCTGATTTGCTGCCGTTGAAGAAATGAGAGTTTCAAGGGCG	3180
Db	3491	AAACGTGATGGAACACGCTCGCTGATTTGCTGCCGTTGAAGAAATGAGAGTTTCAAGGGCG	3550
Qy	3181	CCTCTGAAGAACGAAACAGATTTTCTTTGCTTCCCTTGTGAGCTCTGTCTTAACTGTGGCT	3240
Db	3551	CCTCTGAAGAACGAAACAGATTTTCTTTGCTTCCCTTGTGAGCTCTGTCTTAACTGTGGCT	3610
Qy	3241	TGGATCTGTGTGTTGGTGAACGCTTCTACTGTGTGCTTGGGAGCGCGGAAACCCGGGC	3300
Db	3611	TGGATCTGTGTGTTGGTGAACGCTTCTACTGTGTGCTTGGGAGCGCGGAAACCCGGGC	3670
Qy	3301	AGCCACACACACTCAGCTCTGTGAGGACAAACACCAACAAACGTCGGGAGCAGCTGAAC	3360
Db	3671	AGCCACACACACTCAGCTCTGTGAGGACAAACACCAACAAACGTCGGGAGCAGCTGAAC	3730
Qy	3361	CAGATCAAAAAACCCCATTTGAGAAACATGGGGCCAAACACGGTCCCCATCAAGGATTAACGAG	3420

Db	3731	CAGATCAAAACCCCATTTGAAAACATGGGGCCAAACCGTCCCCCTCAAAGGATTTCGAG	3790
Qy	3421	AACAAGAACTCCAAAATGTCTAAATAAAGGACACACAATTTCTGAAGTAGAAGAGGACGCAC	3480
Db	3791	AACAAGAACTCCAAAATGTCTTAATAAAGGACACACAATTTCTGAAGTAGAAGAGGACGCAC	3850
Qy	3481	ATGGACAAACACACGACGAAAAAGCCCGTTTTGGCAAGCAGCGGCCGTATACGCTGGTGAGAC	3540
Db	3851	ATGGACAAACACACGACGAAAAAGCCCGTTTTGCCAAGCAGCGGCCGTACACGCTGGTGAGAC	3910
Qy	3541	AGAGAAGAGAGCCCCCAACGGCACCGGCACAAACACCACCAACTGGACAAACAAACAG	3600
Db	3911	AGAGAAGAGAGCCCCCAACGGCACCGGCACAAACACCACCAACTGGACAAACAAACAG	3970
Qy	3601	GACAACAGAGACTTGGAAAGTGCCCAGAGCTTAAACCGAATGGAGTACATCGCTATAG	3657
Db	3971	GACAACAGAGACTTGGAAAGTGCCCAGAGCTTAAACCGAATGGAGTACATCGCTATAG	4027

```

RESULT 13
US-09-310-685-3
; Sequence 3, Application US/09310685
; Patent No. 6887475
; GENERAL INFORMATION:
; APPLICANT: Lamb, Jonathan R
; APPLICANT: Dallman, Margaret J
; APPLICANT: Hoyne, Gerard F
; TITLE OF INVENTION: No. 6887475ch
; FILE OF INVENTION: 674525-2001
; CURRENT APPLICATION NUMBER: US/09/310,685
; CURRENT FILING DATE: 1999-05-04
; PRIOR APPLICATION NUMBER: GB 9623236.8
; PRIOR FILING DATE: 1996-11-07
; PRIOR APPLICATION NUMBER: GB-9715674.9
; PRIOR FILING DATE: 1997-07-24
; PRIOR APPLICATION NUMBER: GB 9719350.2
; PRIOR FILING DATE: 1997-09-11
; PRIOR APPLICATION NUMBER: PCT/GB97/03058
; PRIOR FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 6464
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-310-685-3

```

[illegible][illegible]

Db	1751	GAGCCCTCCTGTCGGGAAATTTGGTTAATGGTTATCGCTGTATCTGTGCCACTGCTGCTATGCA	1811
Qy	1441	GGCGATCACTGTGAGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGGT	1500
Db	1811	GGCGATCACTGTGAGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAATGGGGGT	1870
Qy	1501	CACGTGCAGAAATGAATCAACAGAAATCCAGTGTCTGTGTCCCACTGGTTTCTCTGGAAC	1560
Db	1871	CACGTGCAGAAATGAATCAACAGAAATCCAGTGTCTGTGTCCCACTGGTTTCTCTGGAAC	1930
Qy	1561	CTCTGTGAGTGGACATCGAATTAATGTGAGCCTAAATCCCTGCCAGAACGGTGCCTGAGTGC	1620
Db	1931	CTCTGTGAGTGGACATCGAATTAATGTGAGCCTAAATCCCTGCCAGAACGGTGCCTGAGTGC	1990
Qy	1621	TACAACCGTGCAGTGAATTTCTGCAAGTGGCCCGAGGACTATGAGGGCAAGAACTGC	1680
Db	1991	TACAACCGTGCAGTGAATTTCTGCAAGTGGCCCGAGGACTATGAGGGCAAGAACTGC	2050
Qy	1681	TCACAACCTGAAAGACCACTGCCGACGACCCCCCTGTGAAAGTGAATGACAGCTGCAACAGTG	1740
Db	2051	TCACAACCTGAAAGACCACTGCCGACGACCCCCCTGTGAAAGTGAATGACAGCTGCAACAGTG	2110
Qy	1741	GCCATGGCTTCAACGACACACCTGAAGGGGTGGGTATATTTCTCCACAGCTCTGTGGT	1800
Db	2111	GCCATGGCTTCAACGACACACCTGAAGGGGTGGGTATATTTCTCCACAGCTCTGTGGT	2170
Qy	1801	CCTCAGCGGAAGTCAAGAGTCACTGCGGAGGCAAAATCACTGTGTACTGTAAACAAGGC	1860
Db	2171	CCTCAGCGGAAGTCAAGAGTCACTGCGGAGGCAAAATCACTGTGTACTGTAAACAAGGC	2230
Qy	1861	TTCAGGGGAAATATCTGCCATGAAAAATTAATGACTGTGAGAGCAACCTCTGTGAAAC	1920
Db	2231	TTCAGGGGAAATATCTGCCATGAAAAATTAATGACTGTGAGAGCAACCTCTGTGAAAC	2290
Qy	1921	GGTGGCACTTGCAATCGATGTGTCAAATCTCCTAACAGTGCATCTGTGTGACGCTGGAG	1980
Db	2291	GGTGGCACTTGCAATCGATGTGTCAAATCTCCTAACAGTGCATCTGTGTGACGCTGGAG	2350
Qy	1981	GGGGCTTACTGTGAAACCAATTAATGACTGACGAGCCAGAACCCCTGCCACAATGGGGC	2040
Db	2351	GGGGCTTACTGTGAAACCAATTAATGACTGACGAGCCAGAACCCCTGCCACAATGGGGC	2410
Qy	2041	ACGTGTGCGCACTGGTCAATGACTTCTACTGTGACTGTAAAAATGGTGGAAAGAAAG	2100
Db	2411	ACGTGTGCGCACTGGTCAATGACTTCTACTGTGACTGTAAAAATGGTGGAAAGAAAG	2470
Qy	2101	ACCTGCCACTCACGTGACAGTCACTGTGATGAGGCCACGTGCCAACACAGGTGGCACTGCTC	2160
Db	2471	ACCTGCCACTCACGTGACAGTCACTGTGATGAGGCCACGTGCCAACACAGGTGGCACTGCTC	2530
Qy	2161	TATGATGAGGGGGATGCTTTTAACTGACATGTCTCTGCGGCTGGGAAGGAACAACCTGT	2220
Db	2531	TATGATGAGGGGGATGCTTTTAACTGACATGTCTCTGCGGCTGGGAAGGAACAACCTGT	2590
Qy	2221	AACATAGCCCGAAACAGTAGTCTGCTGCCAACCCCTGCCATAATGGGGGCACATGTGTG	2280
Db	2591	AACATAGCCCGAAACAGTAGTCTGCTGCCAACCCCTGCCATAATGGGGGCACATGTGTG	2650
Qy	2281	GTCAAACGCGAGTCTTTTAACTGCTGTGCAAGGAAGCTGGGAGGGGCCCATCTGTGCT	2340
Db	2651	GTCAAACGCGAGTCTTTTAACTGCTGTGCAAGGAAGCTGGGAGGGGCCCATCTGTGCT	2710
Qy	2341	CAGAAATCAATGACTGACGAGCCCTCATCCCTGTTTACAACAGCGGCACTGTGTGGATGGA	2400
Db	2711	CAGAAATCAATGACTGACGAGCCCTCATCCCTGTTTACAACAGCGGCACTGTGTGGATGGA	2770
Qy	2401	GACAACCTGGTACCGGTGCGAATGTGCCCGGGTTTGTCTGGGCCCGCATCTGCAGAAATAAC	2460
Db	2771	GACAACCTGGTACCGGTGCGAATGTGCCCGGGTTTGTCTGGGCCCGCATCTGCAGAAATAAC	2830
Qy	2461	ATCAATGAATGCCAGTCTTCACTTGTGCTCTTTGGAGCGACCTGTGTGGATGAGATCAAT	2520
Db	2831	ATCAATGAATGCCAGTCTTCACTTGTGCTCTTTGGAGCGACCTGTGTGGATGAGATCAAT	2890

[illegible]

Qy	3601	GACAAACAGAGACTTTGGAAAGTCCCGAGAGCTTAAACCGAATGGAGTACATCGTATAG	3657
Db	3971	GACAAACAGAGACTTTGGAAAGTCCCGAGAGCTTAAACCGAATGGAGTACATCGTATAG	4027
RESULT 14			
US-09-199-865-2			
; Sequence 2, Application US/09199865			
; Patent No. 6433138			
; GENERAL INFORMATION:			
; APPLICANT: Zimrin, Ann B.			
; APPLICANT: Maciag, Thomas			
; APPLICANT: Wong, Michael K.K.			
; APPLICANT: Pepper, Michael S.			
; APPLICANT: Montesano, Roberto			
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC METHODS AND COMPOSITIONS			
; TITLE OF INVENTION: BASED ON JAGGED/NOTCH PROTEINS AND NUCLEIC ACIDS			
; FILE REFERENCE: 0036-1U1			
; CURRENT APPLICATION NUMBER: US/09/199,865			
; CURRENT FILING DATE: 1998-11-25			
; EARLIER APPLICATION NUMBER: 60/018,841			
; EARLIER FILING DATE: 1996-05-31			
; EARLIER APPLICATION NUMBER: PCT/US97/09407			
; EARLIER FILING DATE: 1997-05-30			
; NUMBER OF SEQ ID NOS: 16			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 2			
; LENGTH: 5458			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-199-865-2			
Query Match 98.8%; Score 3611.8; DB 3; Length 5458;			
Best Local Similarity 99.8%; Pred. No. 0;			
Matches 3613; Conservative 2; Mismatches 4; Indels 0; Gaps 0;			
Qy	39	CCTAAGCCTCTGTCTCGCCCTGCTCTGTGCCCTCGAGGCCAAGGTGTGTGGGGCTCTCGG	98
Db	9	CCTAAGCCTCTGTCTCGCCCTGCTCTGTGCCCTCGAGGCCAAGGTGTGTGGGGCTCTCGG	68
Qy	99	TCAGTTCGAGTTGGAGATCCTGTCCATGCAAGCTGAACGGGAGCTGCAGAACGGGAA	158
Db	69	TCAGTTCGAGTTGGAGATCCTGTCCATGCAAGCTGAACGGGAGCTGCAGAACGGGAA	128
Qy	159	CTGCTCGGGCGGGCCCGGAAACCGGGAGACCGCAAGTGCACCCCGCAGAGTGTGCAC	218
Db	129	CTGCTCGGGCGGGCCCGGAAACCGGGAGACCGCAAGTGCACCCCGCAGAGTGTGCAC	188
Qy	219	ATATCTCAAAAGTGTGCTCTCAAGGAGTATCAGTCTCCGGTGCACGCCGGGGGCCCTCGAG	278
Db	189	ATATCTCAAAAGTGTGCTCTCAAGGAGTATCAGTCTCCGGTGCACGCCGGGGGCCCTCGAG	248
Qy	279	CTTCGGCTCAGGCTCAGCGCTGTCATCGGGGGCAACACCTTCAACCTCAAGGCCAGCCG	338
Db	249	CTTCGGCTCAGGCTCAGCGCTGTCATCGGGGGCAACACCTTCAACCTCAAGGCCAGCCG	308
Qy	339	CGGCACACGACCGCAACCGCATCGTGTGCTCTTTCAGTCTTCGCCCTGGCCGAGTCTTATAC	398
Db	309	CGGCACACGACCGCAACCGCATCGTGTGCTCTTTCAGTCTTCGCCCTGGCCGAGTCTTATAC	368
Qy	399	GTTCGTTGTGGAGGCGTGGGATTCACAGTAATGACACCGTTCAACCTGACAGTATTATGA	458
Db	369	GTTCGTTGTGGAGGCGTGGGATTCACAGTAATGACACCGTTCAACCTGACAGTATTATGA	428
Qy	459	AAAGGCTTCTCACTCGGGCATGATCAACCCCGCGGAGTGGCAGACGCTGAAGCGAA	518
Db	429	AAAGGCTTCTCACTCGGGCATGATCAACCCCGCGGAGTGGCAGACGCTGAAGCGAA	488
Qy	519	CACGGGCGTTGCCACCTTTCAGATATCAGATCCGGGTGACCTGTGATGACTACTACTATGG	578
Db	489	CACGGGCGTTGCCACCTTTCAGATATCAGATCCGGGTGACCTGTGATGACTACTACTATGG	548
Qy	579	CTTTGGCTGYAATAAGTTCTTCGCGCCCGCCAGCAGATGACTTCTTTTGGACACTATGCTCTGTA	638

; TITLE OF INVENTION: BASED ON JAGGED/NOTCH PROTEINS AND NUCLEIC ACIDS									
; FILE REFERENCE: 0036-101									
; CURRENT APPLICATION NUMBER: US/10/213,329									
; CURRENT FILING DATE: 2002-08-06									
; PRIOR APPLICATION NUMBER: US/09/199,865									
; PRIOR FILING DATE: 1998-11-25									
; PRIOR APPLICATION NUMBER: 60/018,841									
; PRIOR FILING DATE: 1996-05-31									
; PRIOR APPLICATION NUMBER: PCT/US97/09407									
; PRIOR FILING DATE: 1997-05-30									
; NUMBER OF SEQ ID NOS: 16									
; SOFTWARE: PatentIn Ver. 2.0									
; SEQ ID NO 2									
; LENGTH: 5458									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
US-10-213-329-2									
Query Match 98.8%; Score 3611.8; DB 3; Length 5458;									
Best Local Similarity 99.8%; Pred. No. 0;									
Matches 3613; Conservative 2; Mismatches 4; Indels 0; Gaps 0;									
QY	39	CCTAAGCCTCTGCTCGCCCTGCTCTGTGCTGCTGCGAGCCAAAGTGTGTGGGCTCGGG	98						
DB	9	CCTAAGCCTCTGCTCGCCCTGCTCTGTGCTGCTGCGAGCCAAAGTGTGTGGGCTCGGG	68						
QY	99	TCAGTTTCAGGTTGAGATCTCTGCTCATGCAAGACGTGAACGGGAGCTGCAGAACGGGAA	158						
DB	69	TCAGTTTCAGGTTGAGATCTCTGCTCATGCAAGACGTGAACGGGAGCTGCAGAACGGGAA	128						
QY	159	CTGCTGGGGGGGGCCCGGAACCGGGGAGACCGCAAGTGCAACCGCGGACGAGTGTGACAC	218						
DB	129	CTGCTGGGGGGGGCCCGGAACCGGGGAGACCGCAAGTGCAACCGCGGACGAGTGTGACAC	188						
QY	219	ATACTTCAAAGTGTGCTCAAGAGATATAGTCCCGCTCAACCGCGGGGGCCCTGCGAG	278						
DB	189	ATACTTCAAAGTGTGCTCAAGAGATATAGTCCCGCTCAACCGCGGGGGCCCTGCGAG	248						
QY	279	CTTTCGGCTCAGGGTCCACGCTGTCTATCGGGGGCAACACCTTCAACCTCAAGGCCAGCCG	338						
DB	249	CTTTCGGCTCAGGGTCCACGCTGTCTATCGGGGGCAACACCTTCAACCTCAAGGCCAGCCG	308						
QY	339	CGGCAACGACCGCAACCGCAATCGTGTCTTTTCAGTTTCGCTGCGCGGAGTCTTATAC	398						
DB	309	CGGCAACGACCGCAACCGCAATCGTGTCTTTTCAGTTTCGCTGCGCGGAGTCTTATAC	368						
QY	399	GTTCCTTGTGGAGCGTGGGATTCAGATATGACCGTTCAACCTGACAGTATTATGA	458						
DB	369	GTTCCTTGTGGAGCGTGGGATTCAGATATGACCGTTCAACCTGACAGTATTATGA	428						
QY	459	AAAGGCTTCTCACTCGGGCATGATCAACCCAGCCGCGAGTGGCAGACGCTGAAGCAGAA	518						
DB	429	AAAGGCTTCTCACTCGGGCATGATCAACCCAGCCGCGAGTGGCAGACGCTGAAGCAGAA	488						
QY	519	CACGGGGTTGCCACTTTGAGTATCAGATCCGCGTGAACCTGATGATGACTACTATGG	578						
DB	489	CACGGGGTTGCCACTTTGAGTATCAGATCCGCGTGAACCTGATGATGACTACTATGG	548						
QY	579	CTTTGGCTGTAATAAGTCTGCGCGCCCGAGAGATGACTTTTGGACACTATGCTCTGA	638						
DB	549	CTTTGGCTGTAATAAGTCTGCGCGCCCGAGAGATGACTTTTGGACACTATGCTCTGA	608						
QY	639	CCAGAAATGGCAACAAACCTTGCATGGAAGGCTGGATGGGCGCCGGAATGTAAACAGAGCTAT	698						
DB	609	CCAGAAATGGCAACAAACCTTGCATGGAAGGCTGGATGGGCGCCGGAATGTAAACAGAGCTAT	668						
QY	699	TTGCCGACAAAGGTCAGTCTTAAGCATGGGTCTTGCAAACTCCCAAGGTGACTGCAAGTG	758						
DB	669	TTGCCGACAAAGGTCAGTCTTAAGCATGGGTCTTGCAAACTCCCAAGGTGACTGCAAGTG	728						
QY	759	CCAGTAYGGCTGCGAAGGCTGTACTGTGATAGTGCATCCCAACCCCGGATGGCTCCA	818						
DB	729	CCAGTAYGGCTGCGAAGGCTGTACTGTGATAGTGCATCCCAACCCCGGATGGCTCCA	788						

QY	819	CGGCATCTGTAAATGAGCCCTGGCAGTGCCTCTGTGAGACCAACTGGGGCGGCCAGCTCTG	878
DB	789	CGGCATCTGTAAATGAGCCCTGGCAGTGCCTCTGTGAGACCAACTGGGGCGGCCAGCTCTG	848
QY	879	TGACAAAGATCTCAATTAATCTGTGGGACTCATAGCCGTGTCTCAACGGGGGAACTTGTAG	938
DB	849	TGACAAAGATCTCAATTAATCTGTGGGACTCATAGCCGTGTCTCAACGGGGGAACTTGTAG	908
QY	939	CAACACAGGCCCTGCACAAATATCAGTGTCTCTCCCTCGAGGGGTATTTCAGGACCCAACTG	998
DB	909	CAACACAGGCCCTGCACAAATATCAGTGTCTCTCCCTCGAGGGGTATTTCAGGACCCAACTG	968
QY	999	TGAAATTTGTAGCACGCGCTGCTCTCTGATCCCTGTGCACACAGAGGAGCTGTAAAGGA	1058
DB	969	TGAAATTTGTAGCACGCGCTGCTCTCTGATCCCTGTGCACACAGAGGAGCTGTAAAGGA	1028
QY	1059	GACCTCCCTGGGCTTTGAGTGTGAGTGTTCCTCCAGGCTGACCGGCCCCACATGCTCTAC	1118
DB	1029	GACCTCCCTGGGCTTTGAGTGTGAGTGTTCCTCCAGGCTGACCGGCCCCACATGCTCTAC	1088
QY	1119	AAACATTTGATGACTGTTCTCTTAATAACTGTTCCACGGGGGACCTGCGGAGACCTGGT	1178
DB	1089	AAACATTTGATGACTGTTCTCTTAATAACTGTTCCACGGGGGACCTGCGGAGACCTGGT	1148
QY	1179	TAACGGATTTAAGTGTGTGTGCCCCCAGACGTGGGAAACCGTGCAGTTAGATGC	1238
DB	1149	TAACGGATTTAAGTGTGTGTGCCCCCAGACGTGGGAAACCGTGCAGTTAGATGC	1208
QY	1239	AAATGAATGTAGGCCCAAACTTGTGTAAACGCAAACTCTGTAAAGATCTCATTTGCCAG	1298
DB	1209	AAATGAATGTAGGCCCAAACTTGTGTAAACGCAAACTCTGTAAAGATCTCATTTGCCAG	1268
QY	1299	CTACTACTGCGACTGTCTTCCGGCTGGATGGTTCAGAAATTTGACATAAATAATTAATGA	1358
DB	1269	CTACTACTGCGACTGTCTTCCGGCTGGATGGTTCAGAAATTTGACATAAATAATTAATGA	1328
QY	1359	CTGCTTTGCCAGTGTGAGATGACGCTCTCTGTGCGGATTTTGGTTAATGGTTATCGCTG	1418
DB	1329	CTGCTTTGCCAGTGTGAGATGACGCTCTCTGTGCGGATTTTGGTTAATGGTTATCGCTG	1388
QY	1419	TATCTGTCAACCTGGCTATGCGGCGATCATCTGTGAGAGAGACATCGATGAATGTGCCAG	1478
DB	1389	TATCTGTCAACCTGGCTATGCGGCGATCATCTGTGAGAGAGACATCGATGAATGTGCCAG	1448
QY	1479	CAACCCCTGTTGAATGGGGGTCACTGTGAGATGAATCAACAGATTCAGTGTCTGTG	1538
DB	1449	CAACCCCTGTTGAATGGGGGTCACTGTGAGATGAATCAACAGATTCAGTGTCTGTG	1508
QY	1539	TCCCACTGTTTCTCTTGAAACCTCTGTGAGTGGACATCGATTATTGTGAGCCTAATCC	1598
DB	1509	TCCCACTGTTTCTCTTGAAACCTCTGTGAGTGGACATCGATTATTGTGAGCCTAATCC	1568
QY	1599	CTGCCAGAAACGGTGCCAGTGTGTAACACCGTGCAGTGAATAATTTCTGCAAGTGCCCGA	1658
DB	1569	CTGCCAGAAACGGTGCCAGTGTGTAACACCGTGCAGTGAATAATTTCTGCAAGTGCCCGA	1628
QY	1659	GGACTATGAGGCGCAGAACTGTCTCACCTGAAAGACCACTGCGGACGACCCCTGTGA	1718
DB	1629	GGACTATGAGGCGCAGAACTGTCTCACCTGAAAGACCACTGCGGACGACCCCTGTGA	1688
QY	1719	AGTGATTGACAGCTGCACAGTGGCCATGGCTTCCAAACGACACACCTGGAAGGGTGGGGTA	1778
DB	1689	AGTGATTGACAGCTGCACAGTGGCCATGGCTTCCAAACGACACACCTGGAAGGGTGGGGTA	1748
QY	1779	TATTTCTCCAAAGCTGTGGTCTCTCAACGGGAAGTGCAGAGTCACTCGGAGGCAAAAT	1838
DB	1749	TATTTCTCCAAAGCTGTGGTCTCTCAACGGGAAGTGCAGAGTCACTCGGAGGCAAAAT	1808
QY	1839	CACCTGTGACTGTAAACAAAGGCTTCAACGGGAACATCTGCCATGAAATAATTAATGACTG	1898
DB	1809	CACCTGTGACTGTAAACAAAGGCTTCAACGGGAACATCTGCCATGAAATAATTAATGACTG	1868

QY	1899	TGAGAGCAACCCCTGTAGAAA	CGGTGGCACTTGATCGATGCTGTCAACTCCTACAAAGTG	1958
Db	1869	TGAGAGCAACCCCTGTAGAAA	CGGTGGCACTTGATCGATGCTGTCAACTCCTACAAAGTG	1928
QY	1959	CATCTGTAGTACGGCTGGAGGGG	CGCTTACTGTGAAACCAATATTAAATGACTGCGAGCCA	2018
Db	1929	CATCTGTAGTACGGCTGGAGGGG	CGCTTACTGTGAAACCAATATTAAATGACTGCGAGCCA	1988
QY	2019	GAACCCCTGCGACAAATGGGGG	CACTGTGCGGACCTGGTCAATGACTTCTACTGTGACTG	2078
Db	1989	GAACCCCTGCGACAAATGGGGG	CACTGTGCGGACCTGGTCAATGACTTCTACTGTGACTG	2048
QY	2079	TAAAAATGGGTGGAAGAAAGAA	AGCTGCTGCACTCAGGTGACAGTCAAGTGTGATGAGGCCAC	2138
Db	2049	TAAAAATGGGTGGAAGAAAGAA	AGCTGCTGCACTCAGGTGACAGTCAAGTGTGATGAGGCCAC	2108
QY	2139	GTGCAACAAACGGTGGCA	CCCTGCTATGATGAGGGGATGCTTTTAAAGTGCAATGTGTCTGTG	2198
Db	2109	GTGCAACAAACGGTGGCA	CCCTGCTATGATGAGGGGATGCTTTTAAAGTGCAATGTGTCTGTG	2168
QY	2199	CGGCTGGGAAGGAACAACCTGT	TAAACATAGCCCGAACAAGTACTGCTGCGCCCAACCCCTG	2258
Db	2169	CGGCTGGGAAGGAACAACCTGT	TAAACATAGCCCGAACAAGTACTGCTGCGCCCAACCCCTG	2228
QY	2259	CCATAATGGGGGCA	CATGTGTGTTCAACGGCGAGTCTCTTACGTGCGTCTGCAAGGAAGG	2318
Db	2229	CCATAATGGGGGCA	CATGTGTGTTCAACGGCGAGTCTCTTACGTGCGTCTGCAAGGAAGG	2288
QY	2319	CTGGAGGGGCCCACCTGTGCT	CAAGATACCAATGACTGCAGGCCCTCATCCCTGTGTACAA	2378
Db	2289	CTGGAGGGGCCCACCTGTGCT	CAAGATACCAATGACTGCAGGCCCTCATCCCTGTGTACAA	2348
QY	2379	CAGCGGCAACCTGTGTGATGG	AGACAACCTGGTCCGCTGCGAATGTGCCCCGGGTTTGC	2438
Db	2349	CAGCGGCAACCTGTGTGATGG	AGACAACCTGGTCCGCTGCGAATGTGCCCCGGGTTTGC	2408
QY	2439	TGGSCCGAGCTGCAGAAATAA	CAATGAATGACAGTCTTCACTTTGTGCTTTGGAGC	2498
Db	2409	TGGSCCGAGCTGCAGAAATAA	CAATGAATGACAGTCTTCACTTTGTGCTTTGGAGC	2468
QY	2499	GACCTGTGTGATGAGATCAAT	TGGCTACCGGTGTGTCTGCCCTCCAGGGCAACAGTGGTGC	2558
Db	2469	GACCTGTGTGATGAGATCAAT	TGGCTACCGGTGTGTCTGCCCTCCAGGGCAACAGTGGTGC	2528
QY	2559	CAAGTGCAGGAAGTTTCAGG	GAGACCTTGCATCACATGGGGAGTGTGATCCAGATGG	2618
Db	2529	CAAGTGCAGGAAGTTTCAGG	GAGACCTTGCATCACATGGGGAGTGTGATCCAGATGG	2588
QY	2619	GGCCAAATGGGATGATGACT	GTAAATACCTGCGAGTGTGCTGAATGGACGGATCGGCTGCTC	2678
Db	2589	GGCCAAATGGGATGATGACT	GTAAATACCTGCGAGTGTGCTGAATGGACGGATCGGCTGCTC	2648
QY	2679	AAAGGTCTGTGTGGCCCTCG	ACCTTGCCTCCAAGGGGCAACAGGAGTGGCCCGAG	2738
Db	2649	AAAGGTCTGTGTGGCCCTCG	ACCTTGCCTCCAAGGGGCAACAGGAGTGGCCCGAG	2708
QY	2739	CGGCGAGCTGCAATCCCG	CACTTGGAGCAGCAGTGTGCTTCCACCCCTGCATGGTGT	2798
Db	2709	CGGCGAGCTGCAATCCCG	CACTTGGAGCAGCAGTGTGCTTCCACCCCTGCATGGTGT	2768
QY	2799	GGGCGAGTGTGGTCTTCC	AGTCTCCAGCGGTGAAGCAAAAGTGCACCTCTGACTCCTA	2858
Db	2769	GGGCGAGTGTGGTCTTCC	AGTCTCCAGCGGTGAAGCAAAAGTGCACCTCTGACTCCTA	2828
QY	2859	TTACCAGGATAACTGTGCG	AAACATCACATTTTACCTTTAAACAGGAGATGATGTCAACGAGG	2918
Db	2829	TTACCAGGATAACTGTGCG	AAACATCACATTTTACCTTTAAACAGGAGATGATGTCAACGAGG	2888
QY	2919	TCCTTACTACGAGCAACAT	TTGAGTGAATTTTGAATTTTGAAGAAATGTTTC	2978
Db	2889	TCCTTACTACGAGCAACAT	TTGAGTGAATTTTGAATTTTGAAGAAATGTTTC	2948
QY	2979	CGCTGAATATTTCAATCT	ACATCGCTTGGAGCCTTCCCTTTCAAGCAACAATGAATACA	3038

Search completed: February 5, 2006, 12:49:24
Job time : 653 secs

Db	2949	CGCTGAATATTTCAATCT	ACATCGCTTGGAGCCTTCCCTTTCAAGCAACAATGAATACA	3008
QY	3039	TGTGGCCATTTCTGCTGA	AGATATACGGGATGATGGGAACCCGATCAAGGAATCACTGA	3098
Db	3009	TGTGGCCATTTCTGCTGA	AGATATACGGGATGATGGGAACCCGATCAAGGAATCACTGA	3068
QY	3099	CAAAATAATCGATCTTGT	TAGTAAACGTGATGAAACAGCTCGCTGATTTGCTGCGCGTTGC	3158
Db	3069	CAAAATAATCGATCTTGT	TAGTAAACGTGATGAAACAGCTCGCTGATTTGCTGCGCGTTGC	3128
QY	3159	AGAAGTAAAGATTCAGAG	CGGCTCTGAAAGAACAGATTTCTTGTTCCTTGTCTTGTCTTGTCT	3218
Db	3129	AGAAGTAAAGATTCAGAG	CGGCTCTGAAAGAACAGATTTCTTGTTCCTTGTCTTGTCTTGTCT	3188
QY	3219	GAGCTCTGTCTTAACT	TGCTTGGATCTGTTGCTTGGTGGAGCCCTTCTACTGTTGCTCT	3278
Db	3189	GAGCTCTGTCTTAACT	TGCTTGGATCTGTTGCTTGGTGGAGCCCTTCTACTGTTGCTCT	3248
QY	3279	GCGGAAGCGCGGAAGCC	CGGAGCCACACACTCAGCTCAGCTCTTGAGGACAAACCCACAA	3338
Db	3249	GCGGAAGCGCGGAAGCC	CGGAGCCACACACTCAGCTCAGCTCTTGAGGACAAACCCACAA	3308
QY	3339	CAAACGTGCGGGAGCAG	CTGAAACAGATCAAAAAACCCCATTTGAGAAACATGGGGCCCAACAC	3398
Db	3309	CAAACGTGCGGGAGCAG	CTGAAACAGATCAAAAAACCCCATTTGAGAAACATGGGGCCCAACAC	3368
QY	3399	GGTCCCATCAAGGATTA	CGGAAACAGAACTCCAAAAATGTCTAAAAATAGGACACACAA	3458
Db	3369	GGTCCCATCAAGGATTA	CGGAAACAGAACTCCAAAAATGTCTAAAAATAGGACACACAA	3428
QY	3459	TTCTGAAGTAGAAGAG	GACACATGGACAAACACACAGCAAGCCCGTTTGGCAAGCA	3518
Db	3429	TTCTGAAGTAGAAGAG	GACACATGGACAAACACACAGCAAGCCCGTTTGGCAAGCA	3488
QY	3519	GCCGGCGTATACGCT	GCTAGACAGAAAGCCCCCAACGGGACCGCCGACAAACAA	3578
Db	3489	GCCGGCGTATACGCT	GCTAGACAGAAAGCCCCCAACGGGACCGCCGACAAACAA	3548
QY	3579	CCCAAACTGGACAAAC	CAACAGGACACAGAGCTTGGAAAGTGGCCAGAGCTTAAACCG	3638
Db	3549	CCCAAACTGGACAAAC	CAACAGGACACAGAGCTTGGAAAGTGGCCAGAGCTTAAACCG	3608
QY	3639	AATGGAGTACATCGT	TATAG 3657	
Db	3609	AATGGAGTACATCGT	TATAG 3627	